Adaptation on Personalized Public Displays driven by Social Context

Dissertation

an der Fakultät für Angewandte Informatik,
der Universität Augsburg

vorgelegt von M.Sc.

Ekaterina Kurdyukova

Augsburg, 03.03.2014

Erstgutachter: **Prof. Dr. Elisabeth André**

Zweitgutachter: **Prof. Dr. Bernhard Bauer**

Drittgutachter**:** **Prof. Dr. Albrecht Schmidt**

Tag der mündlichen Prüfung: 15.05.2014

Dedication and Acknowledgements

The dissertation is devoted to my grandmother and mother, Valentina Petrova and Olga Petrova, who brought me up the respect to the science and the importance of higher education. Many thanks to my grandmother for showing me how successful a woman can be in a technical research.

Many thanks to my uncle, Jury Morokhovets, for his great contribution to my education, for encouragement to start the dissertation, and for the vivid interest in the research progress.

Many thanks to my main advisor, Prof. Dr. Elisabeth André, for giving me the opportunity to complete the dissertation, for guiding me, for the chance to work in an outstanding lab and conduct the research on international level.

Thanks to the HCM research team of Augsburg University, for the wonderful four years of collaboration. My special thanks to Stephan Hammer and Michael Wissner for their support in Trust-related research.

Special thanks to my further advisor, Prof. Dr. Albrecht Schmidt. The decision to start the doctoral thesis was greatly inspired by his work and the examples how exciting Multimedia research can be. Thanks for the fruitful discussions and encouraging talks; they made an invaluable contribution to my thesis.

Thanks to Prof. Dr. Bernhard Bauer for his readiness to be my second advisor from Augsburg University.

Many thanks to my friends in Munich, Aachen, Moscow, and Milan, for their support, ideas, and active participation in the experiments. Special thanks to Dr. Mohammad Obaid for the wonderful time in Augsburg, for the fruitful collaboration, and the charge of creativity.

Thanks to my colleagues at BMW Group, for their support and encouragement to complete the thesis, in spite of a challenging full-time occupation.

My extraordinary thanks to Dr. Marc Piopiunik, for the unconditional support and motivation throughout the thesis. Thanks for being my example of persistency, research discipline, and an illustration of a brilliant scientist.

Finally, I want to thank Javier San Millan, for his care, wise advices, and the truthful fascination about my research.

Abstract

Public displays gain increasing popularity in research and consumer society. They are wide spread indoors in offices and institutions, and outdoors being seamlessly integrated in the urban scene. They aim to entertain, inform, promote, and persuade.

Sensor technologies brought an important advance in human interaction with public displays. By means of modern ambient and mobile sensors the screen can sense and recognize its spectators. Being able to recognize people standing in front of it, the screen can flexibly adapt its content.

It is a challenging task, however, to recognize the social context, and further, to understand how to adapt the screen. The recognition of context must consider a fusion of comprehensive sources of social context. Adaptation of the screen must consider not only presentation of content, but also interaction with the content. The adaptation must also consider the character of the data presented on the screen, for example, how privacy-critical the content is. The adaptation mechanism must be able to handle a large variety of spectators, appearing as groups or single individuals, having diverse and arbitrary identities.

This thesis investigates adaptation to social context from diverse sides, tackling the described research challenges. First, the author studies how to recognize social context, using ambient and mobile sensors. Being able to recognize the context, the author proposes an approach for automatic learning and real-time adaptation of the screen content. The approach is based on recognition of the spectator groups, and tagging the screen content according to the group's interests. Further, the author studies how social context impacts user preferences in interaction style. The empirical studies take a closer look on mobile, physical, and bodily interaction, analyzing interaction preferences in single- and multi-user scenarios, with single- and multiple displays involved in the interaction process. The thesis treats the question of privacy on personalized public displays. The author investigates which type of social context plays a deciding role in user's choice of the display adaptation in privacy-critical scenarios. To round up the thesis, the author provides an approach for automatic decision making on adaptation behaviour. The approach is based on trust-centered Bayesian Network. Relying on the context data delivered by the sensors, the network generates adaptation decisions, best fitting to the current social context.

The thesis aims to contribute to the knowledge of the researchers and practitioners working in the field of interactive public displays, inspiring them to explore new ways of implicit interaction with the screen – interaction driven by the social context.

Zusammenfassung

Große öffentliche Bildschirme (engl. Public Displays) gewinnen zunehmend an Popularität bei Forschern, Entwicklern und Konsumenten. Solche Bildschirme zielen darauf ab, die Betrachter zu unterhalten, zu informieren, zu überzeugen und zum Kauf zu motivieren

Moderne Sensortechnologien eröffnen neue Möglichkeiten zur Interaktion mit Public Displays. Mithilfe von Umgebungs- und mobilen Sensoren können Betrachter von Bildschirmen erkannt werden. Dadurch lassen sich die Inhalte von Bildschirmen flexibel an den Betrachter anpassen.

Es ist jedoch eine enorme Herausforderung, den sozialen Kontext zu erkennen, zu interpretieren und darauf basierend zu entscheiden, welche Anpassung vorgenommen werden soll. Eine umfassende Erkennung des sozialen Kontexts erfordert es, verschiedenste Datenquellen zu fusionieren. Bei der Anpassung des Bildschirms ist nicht nur die Präsentation der Inhalte zu berücksichtigen, sondern auch die Interaktion mit den Inhalten. Außerdem spielt die Art der dargestellten Daten eine entscheidende Rolle, zum Beispiel bei der Frage, inwiefern die Privatsphäre des Nutzers durch die Darstellung beeinträchtigt wird. Der Anpassungsmechanismus sollte in der Lage sein, die Bedürfnisse einer Vielzahl an Betrachtern zu berücksichtigen, und flexibel zwischen Gruppen und Einzelpersonen mit sehr vielfältigen Identitäten unterscheiden.

Die vorliegende Arbeit befasst sich mit den beschriebenen Herausforderungen und untersucht die Anpassung an den sozialen Kontext aus unterschiedlichen Blickwinkeln. Zunächst untersucht die Autorin, wie sozialer Kontext mit Hilfe von Umgebungs- und mobilen Sensoren erkannt werden kann. Darüber hinaus entwickelt sie einen Ansatz für die Echtzeitanpassung der Bildschirminhalte. Der Ansatz basiert auf der Erkennung von Betrachtergruppen und der Annotation des Inhalts anhand der erkannten Gruppeninteressen. Ferner zeigt die Autorin, welche sozialen Kontexte Nutzerpräferenzen bei der Präsentations- und Interaktionsanpassung beeinflussen. Die Arbeit befasst sich außerdem mit dem Schutz der Privatsphäre bei personalisierten öffentlichen Bildschirmen.

Ein Ansatz zur automatischen Entscheidungsfindung bei der Anpassung rundet die Arbeit ab. Der Ansatz basiert auf einem Verfahren zur Modellierung von Vertrauen anhand Bayesscher Netze. Anhand der von Sensoren gelieferten Kontextdaten bestimmt das Netz die Aktionen zur Anpassung, die am besten zur aktuellen Kontextsituation passen.

Die Arbeit strebt an, mit dem erworbenen Wissen zur Forschung auf dem Gebiet Public Displays beizutragen, und Forscher und Entwickler dazu anzuregen, neue Wege der impliziten Interaktion mit Bildschirmen zu erkunden – Interaktionen ausgelöst durch den sozialen Kontext.

Contents

Chapter 1

Introduction

Public displays rapidly gain popularity and importance in the modern life. They are spread around the urban environments, outdoors and indoors. Information, advertisement or entertainment displays can be seen nowadays in public parks, busy city areas, train stations and airports, at the highways, inside offices, in art exhibitions and museums. They seamlessly integrate in our lives, providing necessary information as soon as we require it.

Public displays target a broad range of audience, from elderly people to couples, teenagers, and children. Their usage target may also vary greatly: from informing people about some events or products to persuading or influencing the audience aiming to change their behaviour. The shape, size, and interaction abilities of public displays can also vary: the smallest public displays include tablet-sized devices; the largest ones include urban displays and interactive walls. A popular medium in indoor environments are vertical displays of a 40-80 inches size and interactive tabletops (see Fig. 1.1).

Sensor technologies open wide possibilities for interaction with public displays. Through the sensors today's displays may better "understand" their audience. Such sensitivity to the audience brings an important technological advance: the interaction between human and the display approximates to human-human interaction. The displays can sense who stands in front of them, what are the needs and interests of these people, how they prefer to interact. Based on this understanding, the displays *adapt* their content: present another topic, change colours or resolution, redistribute the space, hide or mask the displayed data. If several displays are involved in the interaction, some of the data may be moved to another display, for example, to a mobile screen.

Figure 1.1. Public displays in the urban scene. Left: an outdoor public display in the city of Oulu, Finland. Right: an interactive tabletop at an exhibition in Berlin, Germany.

In order to assure the seamless adaptation, the displays must recognize the context in automatic or implicit way. The display should understand its audience without any explicit input from the audience side. By means of ubiquitous sensors the displays can automatically gather information on audience dynamics: who approaches the screen, who leaves, whether people talk to each other or not, etc. Knowing how to react to the changing audience, the displays can seamlessly adapt their content.

1.1 Scenarios of Adaptation on Public Displays

To illustrate the idea of adaptation to social context, imagine the following scenarios:

Scenario 1.

David is a young entrepreneur; he is coming back to his office after a lunch break. David is standing at a tram station in the city center waiting for a tram. At the tram station there is a large interactive touch screen showing the latest news. In order to span the waiting time, David approaches the screen. The screen automatically opens a big window with the choice of the news topics. David is touching the "Sports" section and checks the last results of the football league. A minute later a young couple approaches the screen and stands to the left of David. The screen shrinks a little the "Sports" section of David and opens to the left of it a new window with the choice of all topics. The couple chooses the "International" news. The screen is now split into two parts.

Another person comes to the tram station and stands next to the display. However, he doesn't seem to be interested in the display content: he is consumed by his mobile phone typing a message. The display does not react on his presence: the man doesn't even look at the screen.

Finally, the tram comes. Time ran faster than David thought: he still didn't check when the next football match is going to take place. He heads to the tram… And finds a pop up on his mobile phone: “Would you like to continue reading the Sports section?” Great. The missing news he is reading already inside the tram.

Scenario 2.

In a tourist office of Munich a large public display is installed. Marina and Sergey are tourists from Russia; the young couple has just arrived to the city. It is their first time in the city and they are looking forward to get familiar with the town. They approach the display in the tourist office. The display shows sightseeing tips and proposes some routes to discover the city on foot. Moreover, it shows to the young couple some recommendations on romantic cafes and trendy bars.

Later, a couple of Austrian pensioners approaches the display. The couple lives only one hour of drive from Munich; they usually come to Munich by car to spend here one day. Once the couple comes to the display, the screen shows them the tips on best bakeries in the town as well as recommendations of current public events appropriate for their age category, such as cheese festival or open-air opera.

Finally, a mother with a kid, both tourists from Finland, approaches the public screen. They came to Munich for one week, together with their father who is here on a business trip and works the whole week. Being left on their own, the mother and the kid are looking for any child-friendly entertainment. Once they approach the display, they find a map of Munich leisure parks and current city events for families.

Scenario 3.

Peter works in an editorial office of a magazine. He’s just come back from his holidays in India where he spent two weeks together with his wife. In order to share the Indian impressions with his friend colleague Roger, Peter proposes to watch the pictures at a coffee break.

In the afternoon, Peter and Roger go to the coffee room. In the coffee room of their department a large display is situated. The colleagues use it to discuss some work-related materials. Besides, the display serves for entertainment occasions; for example, the colleagues gather there after work to watch a football match.

Peter loads his pictures to the display and starts the presentation. He is commenting the colourful pictures of Indian temples, rural landscapes, and beach pictures from Goa region. Some pictures show Peter together with his wife, on an elephant, in bikini or in a boat.

Suddenly, a colleague Alex enters the coffee room. He is a good friend of Peter, and he is also familiar with Peter’s wife. Peter invites him to join the presentation. All three colleagues continue watching the pictures.

Few minutes later a stranger enters the coffee room. The person is not familiar to Peter; perhaps, he is a colleague from other department or a visitor of the office. The display automatically switches the picture: it hides the photo with Peter and his wife at the beach, and shows instead a neutral landscape picture with no people.

1.2 Thesis Motivation

In order to assure such an intelligent adaptation, the displays must be able to recognize the social context. Moreover, they have to know how to react on the context changes. The awareness of social context and the ability to react on the context changes approximates interaction between a human and a display to human-human communication. It enables the displays to behave as if they would understand us, feel our presence, sense us [74].

A key to success of such a display is in *implicit* collection of social context. The recognition of social context should not require any conscious input from the spectators. Thus, the users will not interrupt their primary routines, but the display will extract the implicit input from their presence and behaviour. Implicit interaction according to Schmidt et al. is an action performed by the user that is not primarily aimed to interact with a computerized system, but which such a system understands as input [190].

The goal of this thesis is to create an approach for adaptation on public displays driven by social context. First, the approach must consider diverse sources of social context, making sense of context data recognized by multiple sensors. Thus, the display will get understanding of the current situation (who is around, what are the needs of these people) and also trace changes in social context (who is coming, who is leaving). Second, the approach must prescribe how the display should react on the changing social context. In particular, how the content should be adjusted, and how interaction style should be tailored to the current social situation. The main challenge here is to provide both recognition of context and adaptation in automatic way, without asking any input from the audience. Finally, the approach must be able to make decisions on adaptation based on the learned knowledge. Thus, the display will be able to decide autonomously on an adaptation action best fitting to the current social context.

1.3 Research Questions

The research goal of this thesis, described in the previous section, can be decomposed into the following research questions:

RQ1: How to sense the social context?

Sensor data can deliver comprehensive knowledge on social context. The challenge, however, is to understand which of these data is crucial for the current situation. Further challenge considers the fusion and synchronization of data acquired from various context sources. The latter questions are particularly relevant to multi-display environments where several heterogeneous displays are involved in interaction. The social context in such environments may be sensed by ambient sensors, installed in the room, and by mobile sensors carried by the audience.

The thesis aims to design and create a universal approach for social context sensing, facilitated by both ambient and mobile sensors. The approach is scalable to multi-display environments, including the environments with heterogeneous displays (vertical public displays, interactive tabletops, tablets, and mobile devices).

RQ2: How to adapt the display content to the changing social context?

Once the surrounding context can be recognized, the next question arises: how to use the social context to adapt the display? The thesis proposes an approach for automatic adaptation of the display content, based on the learnt interests of the display spectators. The approach describes how to learn the content preferences of the spectators and how to use this gained knowledge to automatically adjust the content in the real-time.

RQ3: How to tailor the interaction with the display to the changing social context?

Possessing the knowledge on the social context, and knowing how to adapt the display content, the system needs to know how people want to interact with the adaptive content. Changing social context is likely to change user preferences in interaction style. For instance, when a person is alone in the room, he or she would rather prefer to interact with a large screen directly by touching its surface. However, surrounded by a crowd of strangers, the person might feel more confident interacting from a distance, controlling the screen, for example, from a mobile device.

The thesis investigates the audience's preferences in interaction styles under various constellations of social context. The author analyzes how social context impacts the audience's choice in interaction way and which factors drive their choice.

RQ4: How adaptation to social context can facilitate privacy protection?

Knowing how to adapt the content and tailor the interaction, the system needs to understand how to handle privacy-critical data. Privacy issue is a typical problem during interaction on public displays, especially when personalized data is exposed on the screen. The awareness of the real-time social context can help the display to choose a protective adaptation strategy, satisfying the privacy needs of all involved spectators. In the thesis the author investigates which particular social context is necessary to know for privacy protection and what is the proper reaction on the context changes.

RQ5: How can a display make automatic decisions on adaptation?

Having an approach for recognition of social context and having defined how to adapt the display content and interaction, the system needs to make the approach work automatically.

The final question considered in the thesis is devoted to automatic decision making process, based on the available social context data. The aim of the approach is to make the display autonomously choose an adaptation decision, best fitting to the current social context.

1.4 Thesis Structure

The thesis is organized in seven chapters. Below a short description of each chapter is given.

Chapter 2 gives an overview of the literature related to the area of public display. First, the author gives a classification of public displays, based on their purpose. Then, the particularities of interaction with public displays are addressed. The author discusses the term context awareness, narrowing the focus to the social context. Further, the chapter gives a deeper insight into the design of adaptation driven by social context. Diverse context sources are considered, as well as various display environments. The author concludes the chapter identifying challenges and gaps in the current state-of-the-art.

Chapter 3 presents an approach for the sensing of the social context. Two frameworks are described in the chapter. The first framework was created using ambient sensors. It can be used for the collection of social context data in a stationary setting, such as a room or a workspace with multiple displays. The second framework was created using mobile sensors. The sensors collect the social context of an individual, by means of the sensors integrated into a mobile phone. Apart from the conceptual design of the frameworks, the chapter provides implementation details, illustrating show cases of the framework usages, and the discussion of advantages and drawbacks of either sensing approach.

Chapter 4 is devoted to adaptation of the display content. The author proposes an approach for automatic content adaptation based on the interests of spectators. The approach includes the training phase, where the display is learning about the interests of the spectators, and the adaptation phase, where the display uses the gained knowledge to adjust the content of the screen. The approach utilizes a fusion of social context data extracted from various ambient sensors.

Chapter 5 investigates how to tailor interaction with the public display, given the knowledge of the social context. In particular, the author explores mobile-based, physical, and bodily interaction techniques. Various interaction scenarios are analyzed: single- and multi-display environments, interaction of single and multiple users. The author studies which interaction techniques are preferred by the users in different contextual scenarios and how social context drives these preferences.

Chapter 6 explores privacy issues which arise when interacting with public displays. The author analyzes the social context data that are critical for privacy protection and shows how to design adaptation driven by these data. The chapter also illustrates the fusion of context data used to identify the necessity of privacy-protective adaptation.

Chapter 7 accumulates the knowledge gained in the previous chapters. It describes an approach for automatic decision making on adaptation actions. The approach is based on probabilistic model of Bayesian network. The network is trained by the data collected empirically, in the public display scenarios with diverse constellations of the social context. The author demonstrates the process of data collection, training, initialization, and validation of the network and shows how the network can be used to generate adaptation decisions.

Chapter 8 concludes the thesis. It summarizes the conducted work, emphasizes the research contributions, and outlines the directions of future research which can be inspired by the topics discussed in the thesis.

Chapter 2

Adaptive Public Displays

This chapter gives a broad insight into the literature and research advances made in the field of adaptive displays. Starting with an overview of the diverse display arts, the author narrows down the focus on context-aware displays, taking a closer look at the displays aware of social context.

2.1 Diversity of Public Displays

Public displays are a powerful medium to spread information, encourage communication within communities, and persuade people in their decisions and behaviour change [157]. Nowadays, the large displays can be encountered in urban environments, public places as train stations or shopping malls, workplaces and offices, and in so called "third" public places, i.e. places shared by certain communities, such as public cafes, restaurants, and other places of public gatherings [136].

The shape, content, and purpose of large public displays vary depending on the location of installation and the target spectator groups [39].

Content design of large public displays is usually inspired by real-life metaphors. Müller and colleagues analyze the design space for interactive public displays [143]. The authors end up with four main metaphors or mental models inspiring the design of public displays: poster, windows, mirror, and overlay. Apart from these general categories, other metaphors are frequently used, for instance, bulletin board [75, 40] or a collage [71]. Moreover, design of a public display can represent an extraordinary digital art composition, featuring a trend of contemporary art [200].

The overview below provides examples of public displays, highlighting their diversity and multi-facet applications in the modern life. Modern public displays are mainly used as information providers, as the means for information exchange, as well as the medium to encourage social interactions and influence the behaviour of spectators.

- *Information Screens*

The large screens placed outdoors can convey information on city sights [94], (see Fig. 2.1), share photo impressions of the citizens [165], or promote upcoming cultural events (Milan).

Figure 2.1. Examples of information screens in city environments. Left: a display in the centre of Milan, Italy. Right: an interactive screen in Copenhagen, Denmark.

Large displays installed in public places are usually used to spread information among the potential spectators, help spectators to find necessary data or just provide an entertainment possibility for a short time span. A good example of all these services can be seen on UbiOulu displays [155]. The network of about 15 touch-displays is installed in the city of Oulu in Finland (see Fig. 2.2). The applications running on the displays vary from pure informative ones, narrating about the history of the city, as well as the city sights and notable places, to recommendations of going out city locations, and just entertainment games.

Figure 2.2. UbiOulu public displays spread in the city of Oulu, Finland.

A large display can convey information reflecting the immediate situation inside the city. For instance, a large screen placed on the street crossing can inform drivers about the current traffic situation. The drivers, thus, can adjust their routes to avoid busy streets.

Such an urban display can also present data reflecting activity of the citizens. For instance, CityPulse display demonstrates the social activity within bars, restaurants, discos, and other going-out locations. The data collected by the citizens' smart phones inform about the density within the city locations, loudness, whether people dance or stand, and how much alcohol is consumed [154].

The large displays placed indoors can also facilitate and enrich community life. For instance, a display placed in a public area of an institution or a university may support information exchange between students [147]. A large display placed at a workspace may facilitate news flow within an office; it can also contribute to community building within the colleagues [75].

A large display installed in a public "third place", such as a bar or a cafe, can facilitate news exchange within the community of its visitors [136]. For instance, a display installed in a cafe in Boston showed to enrich the life of the cafe [40]. The visitors appreciated the informative content, such as announcements of the upcoming concerts and other cultural events. Moreover, the visitors appreciated the opportunity to contribute to the content: they were living short notes, impressions, and artistic scribbles.

Designers explore different ways of data presentation on such informative public displays. The conventional way exploits the traditional text-based news, running lines, or bulletin board inspired notes. Moreover, display information can also be presented in artistic way. For instance, a display presenting a dynamic timetable of local buses may be designed as a digital art piece [200] (see Fig. 2.3).

Figure 2.3. Artistic presentation of bus timetable on a public display, inspired by the style of modern artist Mondrian [200]

The presentation media of public displays does not finish conventional LCD or plasma screens. Designers often experiment with alternative media, such as water flows [203], wearable tissues [116], human crowds, and even buildings [25, 22] (see Fig. 2.4).

Figure 2.4. Blinkenlights project utilizes building surfaces as public displays [22].

- *Facilitating Collaboration in a Work Space*

Public displays showed to be an efficient means to improve communication between colleagues. Thus, a display aware of business or working activity of the team members can help the others to better plan collaborative work and avoid undesired interruptions.

For instance, Huang and colleagues presented a community display aware of the presence and occupation status of the team [91]. A glance at the display enabled the team members to estimate how busy the other persons are, and thus helped them to find a right moment for a conversation.

A display placed in meeting room can facilitate collaboration with remote colleagues. Greenberg and Rounding presented Notification Collage display which served as a window to the meeting room of a distant office [75]. The colleagues using the display could switch their status to visible or invisible, thus opening the "window" to be seen from the other distant office. Such simple metaphor realized through a public display simplified the standard procedure of video conferences; it enriched and eased the communication between the remote colleagues.

- *Encouraging Social Interactions*

Public displays can contribute to community building and encourage interactions between community members. For example, Fass and colleagues presented a display installed in a city cafe [56]. The display was aimed to reflect the content created by community members, such as a group of friends, hobbyists, or residents of a café. Installed in a place visited by the community, the display showed to successfully support memory of common activities of the community members.

A similar installation was done by McCarthy and colleagues [136]. The authors installed a large public display, Community Collage, in a cafe in the University district of Seattle. The display was meant to increase interaction and exchange between community members, the residents of the cafe. Supported by a web-based social network platform, the residents could upload their pictures, videos, or just textual notes which were played back in the cafe in the coming days. The deployment study showed that the community technology such as Community Collage meaningfully improves the sense of community.

Churchill and colleagues describe the deployment of eyeCanvas, a large display installed in a cafe in San Francisco [40]. Unlike the previous example, the eyeCanvas display enabled the residents to leave some information immediately in the cafe, and not from a web platform. During the deployment period, the residents indeed posted numerous notes, funny comments, and artistic scribbles which were later shown on the display. Such a display showed potential to enrich the community spaces and serve as a media for artistic expression of the customers.

The deployments of community-building displays showed to be successful also in the office environments. Xerox's CWall was designed for public information sharing within loosely knit groups and communities within Xerox lab [201]. Plasma Poster Network is another example of a community-building display, installed in Palo Alto Research Center [41]. Both deployments showed positive effect on building community of co-workers.

Cheverst and colleagues equipped university offices with tablet-sized public displays, attached to the doors of the staff offices [38]. The Hermes displays were used by both the office owners and visitors. Being able to receive data via Bluetooth, the displays facilitated information exchange between the university staff and students. Moreover, the touch screen of the displays enabled office owners and visitors exchange hand-written notes and scribbles.

Finally, public displays have potential to enrich the life of rural communities. Cheverst and colleagues installed a display in a community centre of Wray, a small town in UK [204]. The rural users appreciated the Wray Photo Display as a new opportunity to share their impressions, express suggestions, and exchange pictures from the local events.

- *Persuasive Displays*

Finally, the displays may influence and encourage people's behaviour change. Such displays have proved efficient in persuading people to become more physically active. Fish'n'Steps display, installed in a public area of Siemens lab, visualized groups' achievements in daily walking [126]. The ambient display helped members of the group trace their progress, and incited them to compete with each other.

Rogers and colleagues installed a persuasive display at the entrance to the office building, in order to influence people to take stairs and not the elevator [179]. The results of the deployment showed that although people consciously did not notice any behaviour change, they did start to take stairs more frequently.

Ambient public displays have a potential to change people's behaviour towards more environment-aware life style. For instance, Tscheligi and colleagues presented an energy awareness display, integrated into a wall clock [195]. The display indicated the level of energy consumed by the household, and provided recommendations how to save the energy and costs.

Apart from the health and environmental interests, persuasive displays may target commercial interests. Thus, Müller and colleagues deployed outdoor displays offering commercial vouchers in the city shops [145]. The displays mounted into telephone cells persuaded people to profit from a voucher code by realizing it in neighbouring shops.

Dietz et al. presented interactive multi-displays in a retail scenario [49]. The display mode depended on the proximity of the customer and the exposed goods. When the customer was in observation mode, further from the screen, the display was in the idle mode showing some promotional videos. When the customer comes closer to the screen approaching the shelves with the goods, the displays activate to attract the customer: they start a demo promoting nearby located goods. Finally, when the customer approached a certain product very closely or touched the product, the display started a demonstration of the product details.

Another work in the retail field presents a human-like mannequin adaptive to the position of the spectators [140]. The mannequin was installed at a shopping window of a department store aiming to attract passers-by. Equipped with cameras, the mannequin could turn his head in the direction where spectators were detected. Once the spectators were detected, the mannequin started a demonstration of the products in trend. Such a lively public display was appreciated by the potential customers. A solid percentage of the interviewed people confirmed that they've spent more time in front of the shopping window.

2.2 Interaction with Public Displays

Although the majority of large displays around us are still not interactive, one can clearly notice the growing interest to and wide spread of interactive displays. Interactive displays are those which are able to interpret an input from the spectator side and react on this input.

Interaction with public displays can be performed in explicit or implicit way. Explicit interaction presumes a conscious input from the spectator, aimed to trigger an expected display reaction. The spectator, thus, gives the display *a command* to achieve the desired goal. The command can be given via an input device, such as a mouse, a keyboard, or a baton [23]; it can be done though a predefined gesture or speech utterance [108].

Implicit interaction uses as input the actions of the user which are not given as a conscious command. For instance, the display can trace user gestures and analyze how actively the user moves. Depending on the level of user activity, the display adjusts its content. Implicit interaction is often enabled by sensor technologies, such as camera-based detection and recognition mechanisms [211], proximity sensing [76], and microphones [107].

The ways of interaction with public displays can be classified into three techniques: direct, bodily, and mobile interaction.

Direct interaction presumes direct manipulation of the objects displayed on the screen. Such manipulation can be achieved by touching the objects, dragging them with a mouse cursor or a specially tailored pointer, such as a digital pen [175]. Direct interaction technique is usually perceived as intuitive; it exploits a real-life metaphor: in order to manipulate an object, people touch it. The disadvantage of direct technique is the necessity to interact from a short distance: the user has to stand right next to the screen in order to reach the displayed objects. However, many public screens are located far from the users, so that they are not physically reachable.

Bodily interaction is enabled by recognition of user's body proximity, posture, gestures or voice commands. An important factor for the success of bodily interaction is the natural set of the bodily commands which can be mapped to the natural human behaviour. For example, the proximity-based interaction can utilize the metaphor of social proximity zones [211, 143]. The closer the person comes to the display, the more information the person potentially is interested in, and thus, the more detailed information should appear on the screen. Bodily interaction is usually enabled by sensor technologies which scan the display surrounding and interpret the context. Therefore, the display spectators can use bodily technique from some distance. Such distant interaction brings, however, potential inconveniences: the camera-based sensors are often not resistant to occlusions which are probable in public spaces. Microphones used for speech recognition are often not resistant to the noise.

Speech interfaces often exploit a limited set of possible voice commands [108]. Recognition of natural language still remains a challenging problem, especially in a noisy public environment.

Finally, mobile interaction presumes the control of a large screen by means of a specially tailored mobile client. Such a client can map the content of the large screen completely [24], imitating thus the direct interaction with the large screen performed on a miniature mobile screen. Alternatively, mobile device can be used as a pointer [15], serving as a mouse on a large screen. The mobile client can complement the content of the large screen [24, 108, 10], presenting a kind of a control panel. The mobile client may exploit not only the screen of the mobile device, but also the device sensors, such as accelerometer, gyroscope or compass. Since the mobile device is usually connected with the public display wirelessly, the mobile interaction is possible from any distance, independently on obstacles or noise. The main disadvantage of mobile interaction is the frequent necessity to control two heterogeneous screens at the same time; the user's visual attention suffers from the permanent focus switch between tiny mobile display and large public display.

Recent research proposes a mix of direct and mobile interaction on large screens, enabled by NFC displays. Such displays represent matrices of NFC tags. The matrix is either placed directly on a wall, and the display content is projected on it [31], or the tags are attached physically to the digital screen, augmenting its functionality [196]. The user interacts with such a screen by touching the tags with an NFC-enabled mobile device. The touch event, registered by the display computer, triggers the changes on the screen.

The described ways of interaction can be used as a conscious input (command) or as an implicit input. In the latter case, the system analyses the context of interaction, interprets it, and executes a respective reaction. For example, the system can analyse user's proximity to the display and adjust content accordingly: showing more detailed information if the user is coming closer or playing general information if the user keeps distance or just passes by. Such an intelligent ability to sense and interpret the surrounding context is referred to as context-awareness.

2.3 Context-Aware Displays

Understanding context is an important part of informing design [5]. Indeed, the ability to sense and interpret the context brings interactive systems certain benefits. First, the context-aware systems are flexible to distinguish between specific situations. Thus, designers can tailor information and functionality relevant only in specific situations [11, 37]. As a result, the user interface of the system can be simplified, demanding less interaction – and cognitive effort – from the user [45]. Second, tailoring the user interface to the specific context facilitates automation of repetitive or trivial tasks [102]. Third, a system which promptly reacts to context changes can increase security of data and users [170]. Moreover, it can improve the performance of applications in the cases where safety is critical [12]. Finally, system behaviour can be trained, by automatically classifying information according to the sensed context. In a changing contextual situation, such a system can retrieve correct behavioural patterns, thus, supporting and supplementing human memory [117].

The term "context awareness" was introduced by Schilit and Theimer back to 1994 [189]. Since that time a lot of research on context-aware computing has been done [192]. Early works on context-aware computing mostly considered context as location of people and objects [189]. However, in recent works the notion of context applies to a broader collection of factors, such as physical and social environment [138, 52, 28, 3, 1, 191] and activities of users [13].

The term context-awareness is tightly bound with the concept of ubiquitous computing [215]. The continuous sensing of the surrounding context lies in the foundation of the ubiquitous technology, where the always-present devices support the "disappearing technology": they come into play only when the user needs them, or when the user is attending to them [1]. Weiser and Brown characterized this technology as "engaging both the centre and periphery attention" [216, 13]. The context sensing enables seamless interaction and seamless transitions between foreground engaging activity and background peripheral perception.

Dey defined context as "any information that can be used to characterize the situation of an entity" [46]. An entity refers to a person, a place or an object that is "considered relevant to the interaction between a user and an application, including user and application themselves." [46].

Although the definition seems to be complete, it is not specific about which type of information can be used to characterize a situation [102].

Schmidt proposed a model of context which differentiates between the context related to human factors and physical environment [191] (see Fig. 2.5).

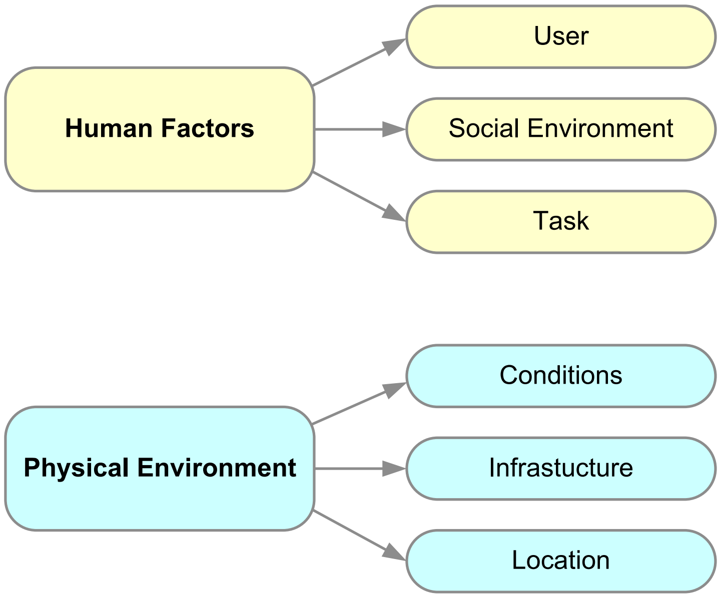


Figure 2.5. Classification of context, according to Schmidt et al. [191].

Human factors describe information about the user, information about the user's social environment, and information about the user's task.

- Information about the user may include user profile, emotional state, or physiological conditions.
- Information about the user's social environment may describe presence of other people, user's belonging to a certain group, as well as group dynamics.
- Information about the user's tasks may reflect the user's current activity, general or intermediate goals.

Context related to *physical environment* describes user location, infrastructure, and physical conditions.

- Location context may reflect user's absolute or relative position, co-location with the other users.
- Infrastructure context may describe available resources, for instance, for communication or for task performance.
- Physical conditions may reflect the level of noise, light, temperature or pressure.

Although the concept of context-aware computing seems fundamentally sound, there are certain challenges which slow down the transition of the idea to useful and usable adaptive systems [162]. First, context interpretation often represents a complex task. The available amount of knowledge on surrounding context is often difficult to interpret not only for machines, but also for humans. Second, it is often unclear how to utilize the knowledge about the context, how to decide which information and functionality to present or to leave out, and how to make use of information already implicitly available to the user [162].

Therefore, there is a need not only in the sensing technologies, but also in a *user model*. The user model enables to reason about the surrounding context, and thus gain understanding how to react to the context changes.

2.4 Social Context

This work investigates public displays adaptive to *social context*. The particular interest to social context is motivated by several arguments. First, public displays are installed in places with high circulation of people (the public). Therefore, it is critical to recognize the current set of spectators, users, and passers-by. Second, according to the media equation [173], people treat computers as persons. Therefore, people are expect that the public display is able to recognize their presence, their behaviour, their interests – and react accordingly, just as another person would do.

Dourish [52, 51] emphasizes the particular importance to social context. In his notion, social context embraces interaction with people, behaviour of people in the environment. Dourish states that context cannot be defined as a stable description of a setting, but arises from activities of people, is sustained by the activities. Social context continually changes in the course of action.

Referring to the model of context proposed by Schmidt and colleagues [191], social context refers to the category human factors. It includes:

Information about the user

- user interests, transferred to the system explicitly, for instance by means of a pre-recorded user profile, or implicitly derived from user behaviour,
- immediate attention, expressed in visual focus on the screen, conversation about the screen content, or a changed behaviour, caused by interruption to observe the screen, user postures and gestures,
- role: involved in interaction (active user) or a passive spectator,
- emotional state, emotional reaction to displayed content,
- physiological data, such as heart beat, pressure, respiration.

Information about the user's social environment

- presence of other people, including number of people at the display,
- constellation of spectators' group, for instance, based on user age or gender,
- social relationships between spectators, for example, friends or strangers,
- disposition of people: proximity to the display and mutual proximity of people,
- conversation between the spectators.

Information about user tasks

- current activity, including user's physical motion or involvement into social activity,
- level of activity, indicating the activity mapped to a certain scale.

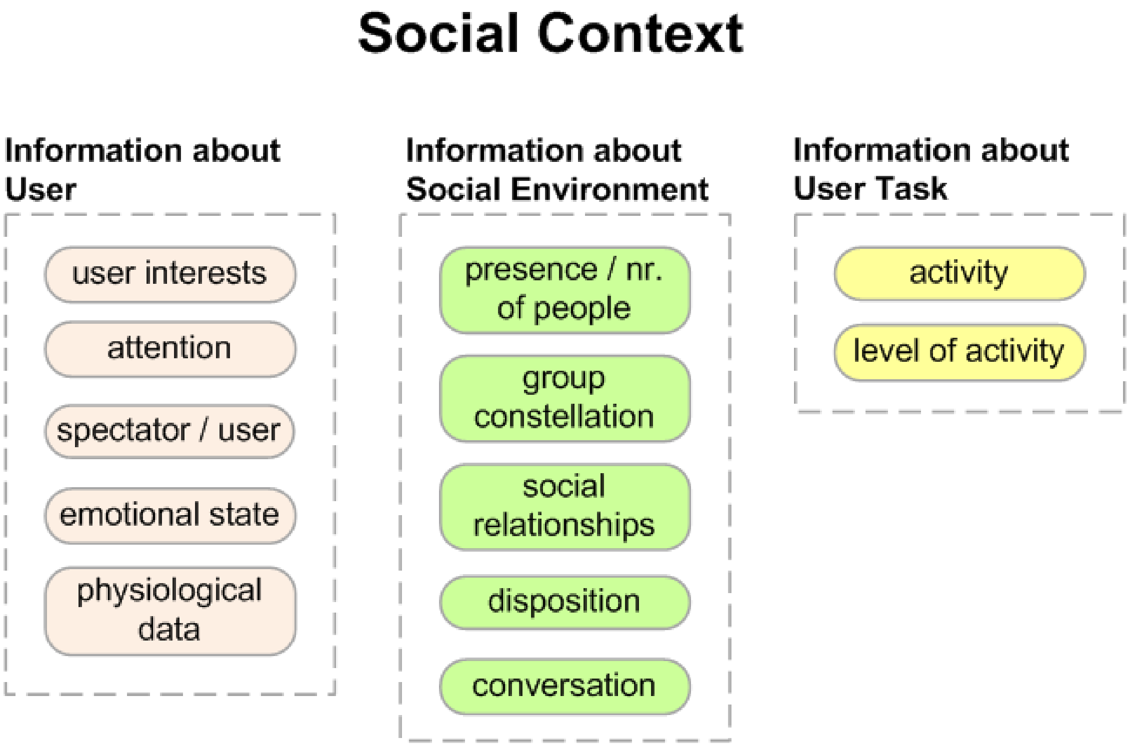


Figure 2.6. Classification of social context used in the thesis, derived from [191].

2.5 Design of Adaptation to Social Context

Plausible interpretation of social context is a necessary ability of interactive public displays. According to the media equation [173], people treat computers as humans. Thus, people interacting with a public display perceive the interaction as a conversation; they expect an intelligent behaviour and reaction from the conversation counterpart – the display.

A human participating in a conversation indeed permanently senses the surrounding social context: who is around, who joins or leaves the conversation, whether the listeners are still interested in the topic, what emotional state they are in, how they react on the given statements. A conversation with a public display follows the same schema: people expect the display to be sensitive to their presence, motion, attention, and emotions. The inability of the display to react according to the social context is likely to provoke people's disappointment, mistrust, and loss of interest.

Therefore, there are two main challenges in designing adaptation to social context.

Challenge 1: Collection of comprehensive context data.

In spite of a great variety of sensors available nowadays, the composition of an exact picture of social context remains challenging. The context data as people perceive it consists of numerous tiny details; they capture not only the primary data, such as presence of co-interactors, conversation, distance to the screen, but also the details as the mood and emotions of the people, people on the periphery, relationships between the people.

Therefore, there is a need in a system which is able to capture as much context data as possible. Thus, such comprehensive recognition will enable plausible interpretation of data and an appropriate display reaction.

Challenge 2: Designing trustworthy adaptation.

Plausible interpretation of context data is a key to the correct choice of adaptation behaviour. However, due to the lack of context data or insufficient learning, the system might choose wrong adaptation behaviour. This mistake can lead to the user disappointment and loss of trust. Therefore, adaptation mechanism should strive for maximization of user trust. Such trustworthy adaptation must assure maintenance of user trust under changing social context.

Below, first the overview of existing context sources is given. The examples illustrate diverse context sources, described in the works on social sciences, and utilized in the works of interaction design, digital art, and real life projects.

After the overview of the works devoted to social context, the necessity in a universal framework for comprehensive recognition of social context is identified. Further, the guidelines for trustworthy adaptation are described; they have to be respected when designing adaptation to social context.

2.5.1 Diversity of Social Context

The following overview provides an analysis of social context sources relevant to the interaction with public displays. The context sources are illustrated by the examples of existing systems, from research or real-life projects. Moreover, the overview discusses context sources which are rarely used in the field of public displays, but rather in other interactive systems. Although these contexts were not explored by designers of public displays so far, they do provide an important source of social context and thus need to be taken into consideration when creating a display sensitive to the social context.

Presence / Number of Spectators

For a display adaptive to social context, the knowledge that some individuals are present in front of it is fundamental. This knowledge is often used to start adaptation: when at least one person approaches the display, the screen switches on [33, starts an animation to attract the person's attention [49] or opens a dialogue [76].

The knowledge about the number of individuals present in front of the screen is often used to divide the display space [76]. The division, however, requires additional information about spatial disposition of the persons: who is located to the right or to the left. So the screen estate will be divided matching the user location.

The awareness of several individuals present in front of the screen can be used not only to split the screen, but also to adjust its content. Villar and colleagues presented a display which detects individuals and their profiles, pre-recorded and stored on a wearable device Pendel [209].

The number of individuals can be also used as an increment, counting how many persons pass by the display to watch its content. For instance, Müller and colleagues counted the number of faces, frontally oriented to the display, which were detected during the day [144]. The number of faces was mapped then to the content of the display at the moment the face was detected. Thus, the authors could estimate the popularity of each content type, counting the faces the content attracted. Although the system provides a realistic estimation of frontal faces, it does not differentiate between single and multiple spectators. Thus, several single spectators and a group of spectators contribute equally to the popularity of the content. This, however, does not correspond to the real observation behaviour.

Finally, the number of individuals can be used to recognize the dynamics of the people within the display area: whether people come or leave. This strategy is especially important for privacy-critical situations. For example, when a new person is entering the room, the display may occlude or hide privacy-sensitive content [211, 181].

Spectator Identity

In a real public scenario, extraction of an individual identity is a difficult and barely feasible task. First, the number of people circulating in public areas such as train stations, shopping malls or galleries, is numerous. Second, their identities are very irregular: many of visitors pass by the display only once.

The general trend in the research on context awareness proposes the shift from identification to personalization [178]. Thus, the unrealistic assumption to capture every single identity is eliminated, giving the way to the more robust recognition of typical user classes or user models.

The projects utilizing the exact identification of individuals rely on face recognition, RFID-based identification or transmission of a Bluetooth profile. For instance, Want and colleagues presented the Active Badge system which was helping office colleagues locate each other [213]. The identification was enabled by identity tags worn by individuals. Ticket to Talk project utilized RFID tags for identification of speakers at a conference [137]. The participants of the conference were supplied by wearable RFID tags which contained their names, professional field and affiliation. Once a participant was having a microphone or just approached a large display, his or her profile appeared on the screen. Such public visit cards eased communication and facilitated professional networking.

Although the identification methods provide a quick way of individual data extraction, they either require an explicit input from the user, during adaptation or beforehand. Preparation of an individual profile and saving it on a storage device requires attention and time, which many users would rather avoid. Face recognition represents a more convenient, ubiquitous approach. The user continues the primary interaction without interruptions for identification procedure. However, face recognition algorithms require an extensive learning phase; moreover, the method is still not robust enough to face occlusions (glasses, hair) and changing light conditions.

Spectator's Interests

A number of research projects rely on individual identity for the sake of interest extraction. The extracted interests are used to adapt the content, of a display or any other medium. Thus, Mahato and colleagues extracted music preferences of cafe visitors, in order to choose the next song to be played [131]. The music preferences were pre-recorded and saved on the visitors' Bluetooth devices.

Villar and colleagues extracted interests of individuals from a wearable device Pendel, in order to adjust the display content to the interests of spectators [209]. If several spectators were present at the screen simultaneously, the interests were merged following a specially tailored algorithm [96].

Similarly, Alt et al. [6] and Karam et al. [101] extracted pre-set profiles of individuals in order to adjust the content of a large public display. The profiles were stored on Bluetooth devices of the spectators. In order to make the profile available for the display, the spectators had to switch on the Bluetooth client manually when being in the proximity of the display.

Although the pre-set profiles definitely represent a reach source of data about the present persons, the manual transmission of the profile is hardly feasible in a real public place. First, people may forget to activate the Bluetooth client. Second, a great number of potential spectators may not possess a Bluetooth client. Thus, these people will not be taken into account when shaping the content – although in reality they do represent the real spectators.

Spectator's Attention

A more ubiquitous way to extract user interests is to trace his or her attention to the content. The term interest here equals to the visual attention, meaning the behaviour of a person watching the content for some period of time, with frontally oriented face.

Having the knowledge of individual attention, the display may dynamically redistribute the content or adjust the content according to the needs of the spectators. For instance, if one of two spectators is not attending to the screen, but rather turned away, the display may enlarge the content of the attentive user.

Visual attention of the spectators also serves as a measure of the content attractiveness. Müller and colleagues used the number of frontally oriented faces to estimate popularity of each content type [144].

The approach can be used the other way around: the display may initiatively attract people's attention when the attention is lost. For instance, a display in a class room may observe the visual attention of the students and play a pre-defined animation when the students' attention is sinking.

Apart from the visual attention, research projects on public display hardly utilize other sources of human attention. However, the works from other interactive fields give more inspiration for design of attention-sensitive displays.

Horvitz and colleagues devoted several projects to attention-sensitive alerting [86]. The authors analyzed utility-directed procedures for mediating the flow of potentially distracting alerts. The goal of the projects was to find a balance of context-sensitive cost of deferring alerts and cost of interruption. Following this goal, the authors investigated user activity and content of notifications.

One of the developed systems, called Priorities, was prioritizing user emails by criticality, choosing modality of notifications about the incoming emails. Each incoming message was assigned a level of criticality, according to its context. The context included the sender personality (colleague, chef, in contacts or not), recipients (only the user or a mailing list), and time criticality. The latter was based on the processing of the message body. The words "soon", "right away", "critical", dates, requests, etc. were contributing to the message criticality.

In the parallel to the message processing, the system measured the person's attention to the display. Thus, the system could estimate whether the current message "deserves" the user distraction.

Another approach to attention recognition is based on eye gaze detection [208]. For instance, Baudisch and colleagues presented a display which shows the content in greater details only in the area of the current user gaze location [14]. Such a method helps to save computational resources for graphically intensive images.

Social Relationships

The importance of relationship as a context is widely studied in social sciences. For instance, Guye-Villeme [78] investigated different aspects of interpersonal relationships. He described the relationships as entities with varying degree of liking, familiarity, trust, and commitment. Rubin and colleagues [182] took a deeper look at the empathic aspects of relationships, emphasizing the degrees of loving and liking in a relationship.

Awareness of interpersonal relationships makes people subconsciously adjust their behavior during an interaction. According to the dynamic boundary regulation model, interacting people constantly adjust variables such as proximity, gaze, and orientation, to the comfortable level, based on their mutual relationships [79, 80].

In the area of interactive displays, the primary interest to the relationships context relates to privacy-critical applications. Indeed, a public displays represents a perfect medium for sharing data in a group. For instance, a group of colleagues may watch pictures from a past corporative event in a recreation room of an office. The pictures, however, might contain a private content not presumed for the eyes of an office visitor. Therefore, the display must quickly recognize appearance of the visitor and promptly protect the privacy-critical data.

The concept of privacy depends on the relationship with the spectator [61, 92]. Similar to the concept introduced by social networks, people need to differentiate between trusted groups [163]. In online network communities, users specify unique policies for the groups of friends, family, colleagues, etc. and thus control the access to their private data [99].

A challenging task, however, is to sense the relationships context. Automatic recognition of the relationships has to rely on a pre-defined relationship structure, such as a social network. The link between the user and the approaching spectator can be retrieved thus by their connection within the structure. A bottleneck here remains a reliable and quick identification of a newly appeared person. Late or incorrect recognition promises a wrong adaptation strategy, and thus user disappointment and mistrust.

Existing public displays have potential to derive the relationships context from identity context. A spectator's identity can recognized either automatically or manually. Automatic identification usually relies on face recognition, whereas manual identification is often based on RFID cards [101] or mobile profiles [6].

Although identification context is often used on adaptive public displays, no works so far were devoted to investigation of relationships as a context.

Group Constellation

Apart from the number of spectators standing in front of the display, it is necessary to distinguish the composition of the spectators' group. The importance of the group composition is known from the social sciences. Mullen, in his work on self-attention theory, described how group composition affects performance and behaviour of an individual [141].

The sensitivity to the group composition context on public display brings a great advantage for adaptation process. For example, a display adapting its content to the interests of spectators can more flexibly distinguish between young or older spectators, couples or single individuals, families or adults.

Although modern technology allows recognition of age, gender, and mutual proximity of individuals, these kinds of context are not explored on interactive public displays. The existing systems usually identify only the number of individuals. Thus, adaptation mechanism does not differentiate who is standing in front of the display: two teenage boys, a couple, or a mother with a kid.

Proximity to the Display

Mutual proximity between people can tell a lot about people's relationships, the character of their conversation, their interests and intentions. This holds not only for people-to-people relations, but also for people-to-objects. For example, a person approaching an object often has an intention to activate or use it.

Proximity in social sciences was studied in detail by Hall [80]. The author introduced the concept of proxemics as "people's cultural perception and use of personal space to mediate their social interactions with others in everyday situations". Proxemics theory studies the interpersonal spatial relationships between individuals and explores how people perceive, interpret, and use micro space around them. Furthermore the theory describes how proximity affects people's interaction and communication with other nearby people.

In proxemics theory Hall correlates physical distance to social distance between people. He categorizes physical distance into four discrete distance zones:

- Intimate: from 0 to 50 cm
- Personal: from 50 cm to 1m
- Social: from 1m to 4m
- Public: from 4m and more

Exact ranges of proxemics zones depend not only on cultural factors, but also on other factors as age, gender, personal relationships [7].

Based on Hall's theory of proxemics, Osmond studied how the spatial arrangement of objects influences people's perception and use of the personal space [161]. The author found out that people indeed perceive different object layouts as sociofugal (separating them from other people) or sociopetal (bringing them together with other people).

Proximity between people located in a same room drives their spatial orientation, eye gaze and orientation [35]. According to the intimacy equilibrium model, people always strive to maintain overall balance towards an optimal proxemic distance. During interaction they constantly adjust the distance; if the distance adjustment is not possible (as for instance in an elevator), people adjust their eye contact and gaze orientation. [35].

Sommer investigated how proximity influences people's preferences in spatial seating arrangements, depending on the task [202]. By means of an experiment, he showed that people do choose different seating locations and relative orientations around a table. Thus, face-to-face arrangements are chosen for competitive tasks, side-by-side or corner-to-corner arrangements – for collaborative conversations.

Proximity context also matters when people choose their location in front of a display. Hawkey and colleagues investigated how people's proximity to a shared display impacts efficiency and enjoyment of collaboration [83]. The authors studied not only proximity between users and the display, but also mutual proximity of users. They found out that collaboration benefits when participants were positioned close together. Interaction with display was felt to be more effective when participants were closer to the display.

The described examples show that proximity is a critical context for interaction not only between people, but also between people and objects.

In the field of interactive public displays, proximity context is a widely explored context source. Most of the proximity-aware displays utilize the context based on pre-defined proximity zones. Designers usually utilize the zones to structure or limit interaction possibilities within each zone.

For example, Funky Wall display introduced by Lucero and colleagues, facilitates collaboration between designer teams and other project stakeholders [129]. The display exploits four ranges of interaction depending on user's proximity: presenting, contemplating, replaying, and exploring. Different functionalities are available within each range. Hand gestures can be performed closely to the screen; they trigger recording of presentations. Gestures performed further from the screen presume exploration; they allow browsing through the presentation.

Vogel and Balakrishnan enable different interaction modalities in different proximity zones [211]. The authors mapped the metaphor of passing-by the screen in order to design interaction zones (see Fig. 2.7).

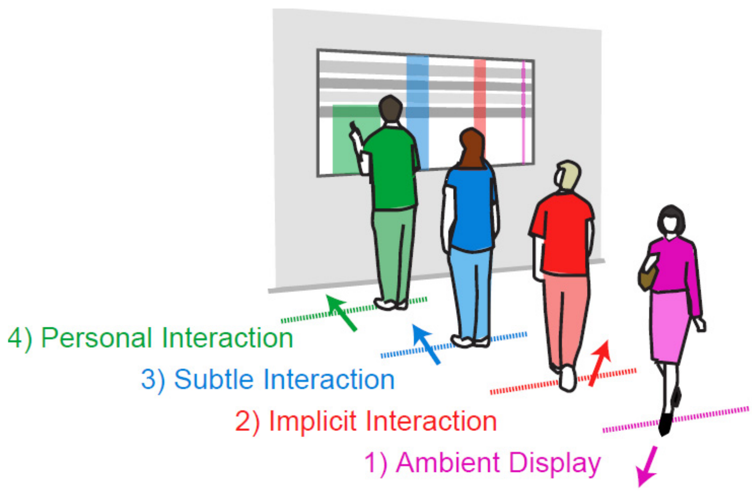


Figure 2.7. Interaction zones defined by the person's proximity to the display [211]

A person passing by the screen and just glancing at the content enables interaction by presence. If the person stops and comes closer to the display, the content is partially personalized; the display enables gestural interaction. If the person approaches the screen in a short distance, the content can be manipulated by finger touch.

Greenberg and colleagues explore a similar concept, mapping proximity to the attention and interest of the passing-by user [74, 133]. Interaction possibilities of the user depend on the proximity zone the user is in. From a further distance, the display can be control by means of a mobile phone or a pointing device. From a shorter distance, the display can be controlled by direct finger touch.

The Range display, introduced by Ju and colleagues follows a similar concept [100]. Proximity zones are used to transition between ambient and authoring modes. Ambient mode allows exploration and browsing of the content. Authoring mode allows edition and modification of content.

An interactive game based on proximity zones was introduced by Finke and colleagues [58]. The authors showed that proximity-based interaction successfully engaged and encouraged spectators to participate in the game. Proximity zone defined the roles of the game participants: bystanders could just glance at the game or enter it, spectators could decode and observe the game, actors could make input, get feedback, and obtain results.

Apart from strictly defined zones, research projects often utilize continuous proximity sensing. For example, Harrison and colleagues presented a display adjusting its zoom level according to the user proximity [82]. The closer the user leans towards the PC display, the larger the content is zoomed. Such semantic zooming helped users avoid dramatic leaning towards a PC; it increased user performance and comfort.

Proximity to the display often defines the role of the users. Bringull and Rogers studied people's activity patterns around a large public display [29]. They found out three roles of passing-by users: engaged in direct interaction, bystanders whose activities are related to focal awareness of the display, and bystanders with peripheral awareness to the display. The authors propose an approach to motivate individuals to interact with the display. The display applications have to support transitions between the role boundaries.

Social Activity

A public display reflecting a level of social activity of an individual within a group can regulate communication within a group. Thus, Ambient Suite display by Fujita et al. gives a feedback on how actively an individual contributes to the group conversation [64]. The feedback is presented as a circle projected on the floor around the individual's feet (see Fig. 2.8). The larger the circle, the more actively the individual contributes. The authors showed that such a display has a potential to improve the conversations of unfamiliar people at a party.

Figure 2.8. Ambient Suite: people interacting in an ambient room [64].

Apart from individual social activity, Ambient Suite display was able to sense the total level of social activity within the group. When the group activity was low, the display on the wall started to show content related to the interests of the group, such as football or art pieces. The topics of interest were recorded in advance, based on the interviews with the group participants. The level of social activity was sensed based on individual's verbal activity, gesticulation with the hands, and head shaking.

Emotional State

Public displays may reflect the emotional state of the audience, in order to encourage or just entertain the spectators. Exeler and colleagues placed a display in a recreation room of an office which showed a face reflecting the emotions of the audience [54]. The goal of the face was to increase the social value of the display: the face expressing the same emotions as the spectators was aimed to show participation in the mood of the room. The recognition of spectators' emotions was based on individual face expressions. The system however could reflect emotions of single faces. It could not generate a merged emotion resulting from two different emotions recognized on the faces of two spectators.

Other examples of emotion-aware public displays can be found in the field of digital arts and theatre. Galaxies project presented a projection-based public display reflecting the emotions of the passers-by [95]. The projection on a wall of a public building represented a group of stars, a galaxy, which was following passing by people. Emotions of the passers-by recognized from their voice, defined the appearance of the galaxies.

An emotionally responsive augmented reality installation, Emotional Tree, represented a visualization of a tree on a public display [65]. The tree reacted on emotions of the spectators by changing its colours and vital appearance.

PuppetWall represents a digital theatre installation where users control puppets on a large display by means of batons specially tailored batons [123]. The emotion of the users extracted from their voice is used to dynamically adjust the scene. If the conversation of the users reaches negative and aggressive tone, the sky of the scene becomes cloudy, darker, and a storm begins.

Physiological Data

Public displays aware of physiological signals incoming from spectators can be encountered in art projects. For instance, Pulse Room installation utilizes an ambient display to transmit the sentimental timeless presence of exposition visitors [167]. The ambient display was represented as a matrix of light bulbs attached to the ceiling of a large exposition room. By touching a sensor installed in the room, a visitor could light one of the bulbs. The bulb started to flash in unison to the visitors heart beat. The entire room therefore was illuminated by the bulbs flashing in different rhythms, repeating the heart beats of the visitors (see Fig. 2.9).

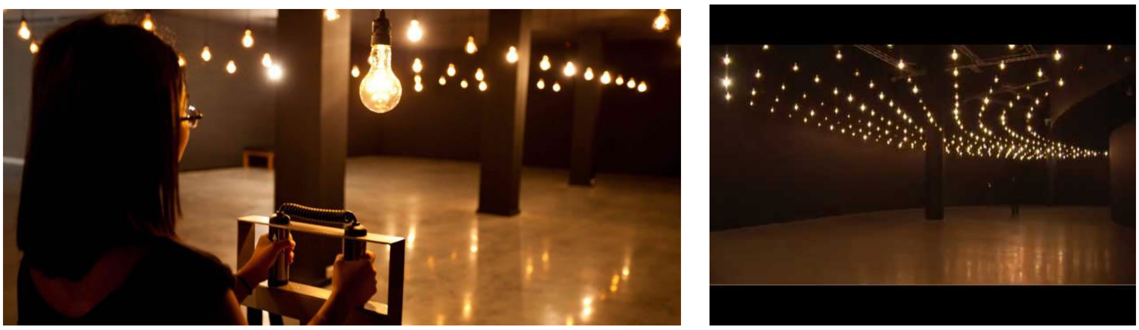


Figure 2.9. Pulse Room [167]

HeartChamber Orchestra is an artistic project where heart beat of musicians is rendered at a large display installed at the concert stage [84]. Visualization and animation of the display are triggered by the level and changes of the musicians currently performing on the stage.

To summarize, modern interactive displays utilize diverse kinds of data for adaptation; however, each of them focus on a certain type of the context.

In order to assure a sufficient picture of the social context, a public display needs to capture comprehensive data, covering as many sources as possible. Therefore, there is a need in a framework which is capable to recognize a vast variety of social context.

The frameworks provided so far by Greenberg et al. [76, 132], Müller et al. [144], and Fujita et al. [64] do offer the recognition of various types of social context. However, none of them covers all the spectrum of context data which is possible to recognize with the modern sensor technology (see Table 2.1).

	Greenberg	Müller	Alt	Fujita	**Kurdyukova**
Number of Spectators	v	v	v	v	**v**
Spectator Identity	v		v	v	**v**
Spectator Interests		v	v		**v**
Spectator Attention	v	v	v		**v**
Social Relationships					**v**
Group Constellation			v	v	**v**
Proximity to Display	v				**v**
Social Activity	v			v	**v**
Emotional State		v			**v**
Physiological Data					**v**

Table 2.1. Comparison of social context data recognized by research frameworks.

The objective of this work is to provide a framework for recognition of social context, using the contemporary sensing mechanisms and covering a wide range of context sources. The framework aims to provide recognition mechanisms capturing the data in implicit way. Thus, the spectators will not have to execute any addition actions in order to help the display collect the data on the social context. The additional actions, such as activation of a Bluetooth client, are rather improbable and distractive in a scenario of a public display usage. Therefore, the mechanism for capturing the social context should work independently on user input.

2.5.2 Designing Trustworthy Adaptation

Adaptation on public displays brings certain advantages and risks. Due to the implicit nature of adaptation, the users often miss the causality behind the adaptive behaviour. Moreover, a high degree of autonomy in adaptive displays may leave the users with the feeling of control loss. Limited amount of transparency and controllability leads to the loss of user trust. As a result, the users feel insecure, frustrated, and are likely to abandon the system. The research goal of this work is to optimize the system actions in a ubiquitous display environment, in order make adaptation design transparent, controllable, and thus trustworthy.

Research on trustworthy design indicates transparency and controllability as the main aspects supporting user trust. Glass and colleagues emphasized the importance of transparency and control in the design of trustable agents [69]. Graham and Cheverst studied interaction paradigms that maintain trust in mobile guides [70]. The authors identified that the lack of transparency and control potentially diminishes user trust. Cheverst and colleagues designed a system that dynamically adjusts to a learnt user model [36]. The authors emphasize the importance of sufficient transparency and comprehensibility of the system and the need to control the existing user model. Bellotti and Edwards [18] as well as Lim and Dey [124, 125] claim that intelligibility significantly improves user trust in context-aware systems. If private data is involved in the adaptation, transparency and control gain even a greater importance. Langheinrich claims that ubiquitous systems should explicitly inform users of aspects that relate to their privacy [118, 119]. The users should be empowered to cancel unauthorized actions. De Vries and colleagues stated that control is positively correlated to user trust [48]: the more control in the hands of users, the more trust the users perceive. However, Schmidt-Belz showed that user control is important rather on semantically high level, related to the system functionality and logic [193]. The control over low-level technical details, such as formats or device capabilities is preferred to be overtaken to the system.

Despite the evident importance of transparency and controllability, no research has been done so far to investigate the guidelines for trustworthy adaptation on public displays, i.e. how and to which extent transparency and control should be provided. An extreme level of transparency or controllability can significantly burden the interaction and cause unacceptable distractions. Therefore, the research should strive for a design trade-off that affords high level of user trust, but keeps interaction comfortable.

The research goal of this work comprises the optimization of the system actions to reflect adaptation in trustworthy and comfortable way.

2.6 Summary

Context-aware public displays possess intelligence to react to the changing in the surrounding context. Such sensitivity enables them to accommodate to the changing situation promptly and according to the expectations of the users.

Sensitivity to social context is particularly valuable for public displays. According to the media equation [173], people treat computers as persons. Thus, the users expect from the display equal abilities to perceive and interpret the surrounding social context.

To reach the level of these expectations, displays meet several challenges. First, they should be able to recognize a vast amount of heterogeneous context data. Indeed, social context varies from presence and number of spectators, through the age, gender, and relations between spectators, and to the spectators' emotions and physiological data.

Although modern sensor technologies enable to recognize the vast types of context data, no works so far aimed to create a framework for recognition of social context. One objective of this work is to provide the framework for the recognition of a wide spectrum of social context data.

Automatic adaptation to the changing social context brings a risk to be misunderstood or misinterpreted by the users. Moreover, the undertaken adaptation can deviate from the user expectations. Therefore, a challenge for creators of adaptive displays is to design adaptation in trustworthy way, assuring sufficient transparency and user control, preserving however comfortable interaction. Another objective of this work is to investigate the design aspects for trustworthy adaptation, and propose a mechanism for trust maintenance, even under uncertainty of available social context data.

Chapter 3

Recognition of Social Context

Awareness of social context is a crucial ability for an interactive public display. Modern sensor technologies enable to approximate the recognition of the surrounding context to the human sensing. The sensors can collect social context data in two ways: they can either be integrated into the room or they can be worn by individuals. The first kind of sensors, ambient sensors, are installed in the area where the context is collected. The sensors can be mounted on the walls, ceiling, or any other fixed position in the area around a display. Such sensors deliver the context data which are relevant only to the current room. The second kind of sensors are attached to the humans; they are carried by the individuals and collect context information independently on the person's location.

Both ways of data collection have their advantages and drawbacks. The stationary sensors provide the entire picture of the present context: who is around the display, and how the situation changes over time. However, the sensors are bound to a specific location and thus the area of their coverage is not scalable and always limited. Moreover, they cannot trace the history of each individual spectator, providing information only related to the individual's actions in the room. Mobile portable sensors, on the contrary, can trace the whole history and dislocations of the people; they collect the data independently on the place and thus can be used, for instance, to trace the social context in a larger areas, such as entire cities. However, in order to have realistic data on social context, the sensors have to be attached to each individual. This requirement is still hard to assure in the modern society.

In this chapter I present two frameworks for recognition of social context: the Ambient Sensing Framework, based on ambient sensors, and the Mobile Sensing Framework, based on mobile sensors.

3.1 Sensing the Social Context with Ambient Sensors

The main challenges for a framework of ambient sensors are to deliver precise data on the *current social context* and to trace how the *context changes* over time. The range and the detail level of the context can vary greatly. The essential context data comprise of the presence of the spectators, their number, and their dynamics: who came in, who left. This set can be extended providing more and more detailed information about the present people: how close people stand to the display, whether they talk or not, what exactly they are discussing, whether people are familiar or not, which age and gender they have, etc. Having this information the display may undertake a certain action, for example, occlude a content on the screen, redistribute the screen space or move the content to another display.

Indeed, in a room where multiple displays are present, the availability of a display is an important context. The decisions on adaptation of content can depend strongly on which alternative displays are available, whether they have private or public character, and whether they potentially possess the user attention.

User attention is the ultimate context neccesary for generation of adaptation decision. This context can be derived from other available context sources. For example, attention to a tabletop display can be derived from user's touch behaviour, from user proximity to the table. Attention in this case equals the potential that the user will notice the content on the display.

To summarize, a framework of ambient sensors needs to provide real-time data on the current social context and enable to trace the context changes. The data should be as precise as possible and should deliver as complete context picture as possible. Moreover, if the ambient room contains multiple displays, the knowledge about the availability of the displays is necessary.

3.1.1 Existing Approaches to Ambient Context Recognition

The essential task for a display aware of social context is to detect the presence of spectators. Detection of spectators can be done either automatically or based on user initiative.

The latter case refers to the systems where the users identify themselves explicitly, for instance, using RFID cards [137] or by activating a Bluetooth client [6]. The advantage of such an approach is the precise picture about the present spectators. The identification mechanisms, such as RFID cards or Bluetooth profiles often contain additional data about the users, for example, their music tastes [131], fields of interest [6] or social network [107, 109]. Therefore, the displays get quick access to valuable information and, hence, can even better tailor the adaptation. The big disadvantage of the explicit identification is the likelihood that the users forget to identify themselves. Moreover, some spectators in the public place may simply not possess the means for identfication and thus be excluded from the collected social context. Such scenario, however, can easily lead to a wrong adaptation decision, since a part of real social context is missing.

A more promising approach for context recognition in public places is implicit sensing. The implicit sensing is supported by sensors integrated into the room, such as cameras, microphones, and motion trackers. The sensors constantly scan the area of the display and update information about the scanned context. The spectators here focus only on their main activities and are not required to do any input.

Modern cameras allow recognition not only of human presence, but also go deep into the face analysis. For example, Müller and colleagues attached cameras to public displays to analyze the amount of people around the display [144] as well as emotions of each person [54]. Emotion recognition was based on facial expressions identified on each detected face. The approach can be further extended to face recognition, i.e. identification of the owner of the detected face. Such identification, however, is hardly possible in a large scale, in a public place with many changing individuals. Face recognition requires a prior training. Thus, it can be used on a limited constant set of spectators, such as a community group or working colleagues.

Motion capturing systems offer an advanced approach to the recognition of social context. In order to enable the recognition, spectators are wearing special markers. The markers are attached to a fixed position on the human body, for example, to a cap or fingers. Cameras installed in the ambient room trace the position of the markers and thus deliver a very precise information about the location and movements of the persons. If the markers are containing a unique identifier, the cameras can also recognize the identity of the tracked person.

Using a motion tracking system Vicon [76], Vogel and Balakrishnan constructed the interaction framework for recognition of social context in front of a large vertical display [211]. The authors equipped the test team with the caps containing the markers. The framework was able to recognize the identity of the user, the distance between the user and the display, orientation of the user's head, and the speed of the movement.

Another framework based on the Vicon motion capturing was presented by Greenberg and colleagues [50]. The framework called Proximity toolkit delivered highly accurate data on user proximity to the display, user orientation, and movement. The framework enables to model proxemity based interaction, by definition of series of easy-to-program events, objects in the space, and possible intersections of the objects with the trajectory of the user motion or orientation of user sight. The model allows to specify which adaptation should be performed on the display, should the user cross the specified proximity treshold.

The show case of the Proximity toolkit was presented in ViconFace project [50]. ViconFace was a comic face rendered on a large vertical display. Depending on user orientation, proximity, and touch events the face changed its mimics (see Fig. 3.1).

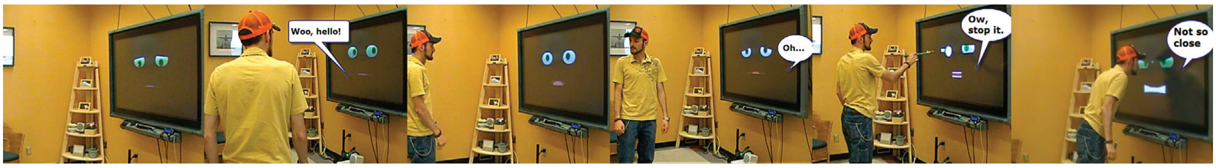


Figure 3.1. ViconFace: reacting on user proximity, orientation, and touch events [50].

The main disadvantage of the motion caputing approach is the necessity to wear the markers. The experiments demonstrating the abilities of motion capturing are usually conducted in lab environemtns. In a real public area it is, however, unprobable that the regular spectators will be one day equipped with the necessary markers.

The presented frameworks [144, 211, 50] are suited for adaptation of a single public display. However, an ambient room can be equipped with several displays. For instance, a modern project room can be equipped with a large vertical display and an interactive table. Moreover, nowadays a popular multidisplay environment represents the combination of a large public display and a small private display, such as a mobile phone or a tablet PC.

Adaptation on multiple displays should consider not only the social context, but also the *display context*. The display context reflects the availablity of the displays. Moreover, the system should know which displays can potentially have *user attention*, and thus can be involved into adaptation.

A framework utilizing the displays context and potential user attention was presented by Elting [53]. The author used an artificial intelligence mechanism to distribute and migrate information according to the movement of the user. The framework takes into consideration the size of the displays and the possible output modality. If the user watches the news in a living room with a large display, the information is given as video; when the user relocated to the kitchen where only a radio is available, the news are continued to be transmitted, but only in audio. Although the system flexibly utilizes the display context, it treats the social context in very minimized way: the only context considered in the framework is the user presence.

Persona UI Framework is another system utilizing social and device context in a multi-device scenario [205]. The authors define the device context in input and output modalities of the available devices. By means of the sensors distributed in an ambient room, the framework orchestrates the migration of content from one display to another, according to the changing social context. The social context is represented by recognized identities of the persons, their speech and gestural input.

Gaia OS framework provides context-sensitive data distribution within heterogeneous displays [180]. The framework recognized device context by means of mobile sensors. Moreover, it gathers social context: presence and identity of users, their orientation, and the services they use. The identity is retrieved by means of RFID tags, markers or IR sensors. The framework enables, for example, automatic saving of a document opened on a large screen when the user leaves the screen. The document is saved locally, and sent to the mobile device of the user.

Context Toolkit, presented by Dey and colleagues [188] describes theoretic approach for collection and spreading of context data, inclusing social context. By means of the sensors integrated into the environement, the framework delivers the data on present users, their identity, and activity. Further extension of the framework also enabled recognition of device context [47]. The framework describes the devices communicating between each other and exchanging the context data. Although the framework describes a powerful means of context data collection, it has rather a theoretical character. The details of implementation and the showcased were not presented so far.

To summarize, there is a number of frameworks enabling collection of social context. Some frameworks also provide recognition of device context, critical for adaptation in a multi-device environment. The existing frameworks ususally focus on specific and limited set of social context. However, in order to approximate to the human sensing of the social context, a framework needs to collect as much data as possible. Moreover, the framework must be able to collect the context in a multi-device enviroment containing heterogeneous displays, such as public displays and private mobile phones.

The objective of this work is to build a universal framework for recognition of social context, the Ambient Sensing Framework. The framework aims to inform designers in possible methods to collect the context data and show how to interpret the acquired context information. Thus, the framework can be reused for various application purposes and flexibly extended for specific needs.

3.1.2 Ambient Sensing Framework

The goal of the framework is to process multiple sources of social context, as many as modern ambient sensors allow, and provide as an output the interpreted data in a simple and understandable form. Thus, designers of adaptive displays can easily use the framework choosing the context data relevant to their project.

In order to be functional in a multi-display environment, the framework was developed for a setting containing a large public display, an interactive tabletop, a tablet PC, and a mobile device. These devices cover the range of ubiquitous displays described in the vision of Weiser [215].

The multi-display environment is equipped with diverse sensors: cameras, microphones, touch sensors, and accelerometers. The cameras are fixed on the walls or large display frame. They perform face detection and face recognition. The microphones are also fixed in the room, next to the static displays: large screen and tabletop. They detect the volume in the area and record speech utterances. The touch sensors on static and mobile displays detect the fact of touch interaction. Accelerometers integrated into mobile devices detect the motion of the device.

Figure 3.2 illustrates the data processing within the Ambient Sensing Framework. The processing is performed in three steps:

1) Raw data is collected by the sensors: cameras and microphones mounted in the room, touch surfaces and accelerometers belonging to the displays. Parallel to the sensor data, the signals from available displays are sent to the framework.

2) The raw data is processed into the material relevant to social context recognition. Faces are extracted from the camera image. Volume and speech utterances are extracted from the volume data. Touch events are detected on the touch surfaces. Motion is detected on the mobile devices.

3) The processed data arrives to the interpretation block. Here the detailed information is transformed into social context signals. For instance, every detected face is analyzed for the proximity to the display, gender, age, and emotions. The concrete interpretation of each context signal is given below.

The device context delivers information on the currently available displays and the displays having user attention.

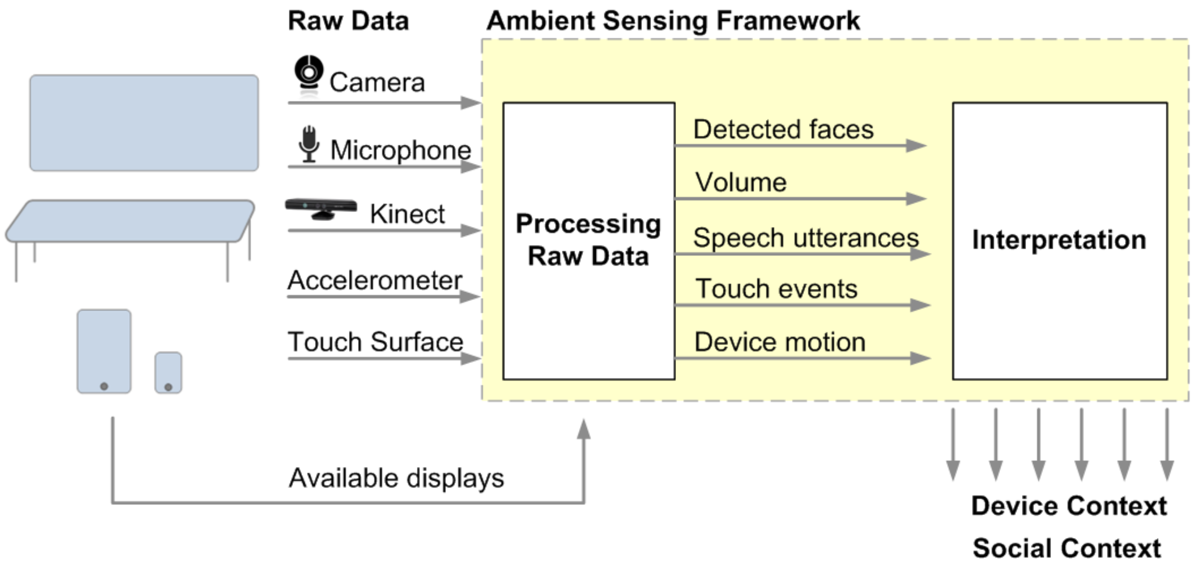


Figure 3.2. Schematic representation of Ambient Sensing Framework.

Framework collects raw data every 0.5 seconds. The processing is performed immeiately, with the same frequency. However, the Raw Data Processing block does not always send the processed data to the Interpretation block. First, the processed data is compared to the data which was sent to Interpretation block the last time. If the data contain some deviations, it means that the social context in the room has changed. For example, there are more detected faces than in the last sent state.

The update in the social context data is generated only if the processed data deviates from the last sent state. Therefore, it is critical to update the current output of the social context, and thus necessary to send the fresh data to the interpreter.

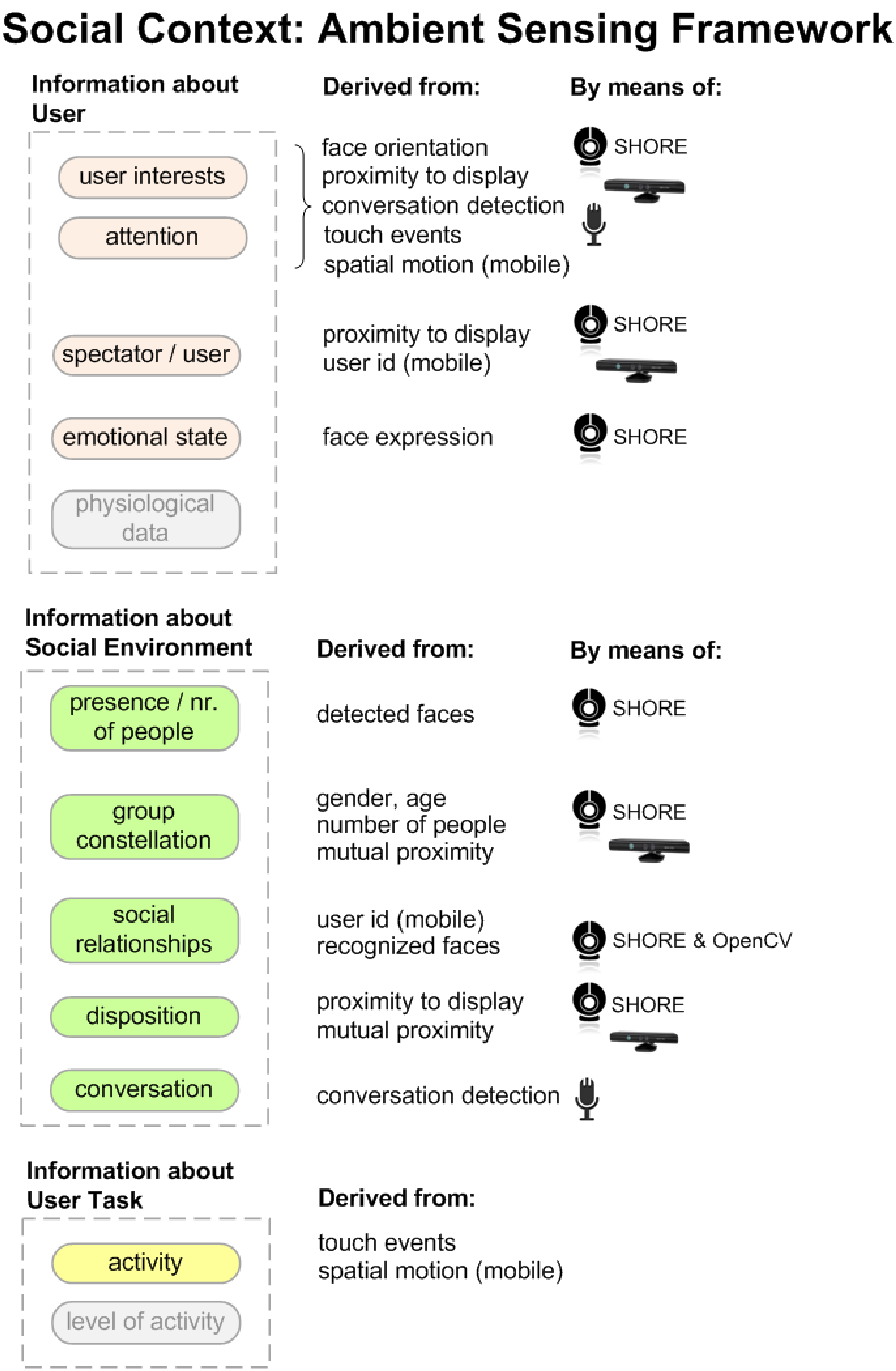


Figure 3.3. Social Context recognized by the Ambient Sensing Framework.

If there is no deviations in the processed data compared to the last state, the data is not sent to the Interpreter. No changes in social context are detected and thus there is no need to update the social context.

3.1.3 Context Sensing

The framework was implemented for a multi-display environment containing various types of the displays: public displays, tabletops, tablet PCs, and mobile devices.

For the experiments conducted within this thesis, the following setting was used. As public display the non-touchable Samsung SyncMaster 550EX was taken. The screen has the size of 55 inches in diagonal; it was connected to a computer running Windows 7 OS. A Genuis web camera with integrated microphone was attached to the bottom frame of the display. As the tabletop the Microsoft Surface was used; the table's computer runs Windows Vista OS. As the tablet PC the Asus Slate EP 121 was taken; the tablet runs Windows 7 OS. Finally, as the mobile device the Samsung Nexus was taken; the device runs Adroid OS.

Figure 3.3 summarizes the context recognized by the Ambient Sensing Framework, giving details on the sources the context is derived from and the sensors involved into the context collection. The following description explains how the framework recognizes distinct context data in the multi-display environment.

Presence of People / Number of People

The number of present people is derived from the number of faces detected by the camera. Face detection mechanism is supported by SHORE software [114] integrated into Social Signal Interpretation (SSI) framework [214]. SHORE software provides real-time face detection; it also recognizes human emotions based on the facial expressions, and estimates the age of the detected face. The SSI framework is a software tool for real-time recording, analyzing, and recognizing of human behaviour. It is able to recognize gestures, mimics, gead nodes, and emotional speech.

The face detection is based on pattern recognition. The system analyzes the camera picture and extracts the areas which have a similar pattern to the human eye brows, nose, and lips. The camera picture is analyzed every 0.3 sec. As an output the system delivers a vector containing of unique user IDs. For each ID the system provides four coordinates: the coordinates of a rectangular outlining the face. The IDs are given only to the detected faces; if no face is detected, the system delives an empty string.

Figure 3.4 left, illustrates the example output of the system. In testing mode, the camera image is overlaid with the detection layer: the outline rectangular and the location of person's eyes.

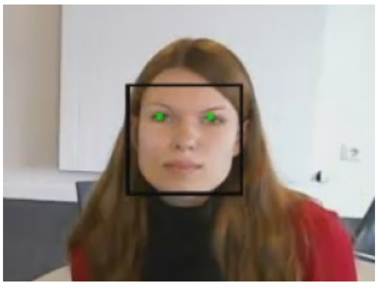


Figure 3.4. Output of SHORE software: correctly detected face (left) and false positives (right).

Important to notice, the pattern-based face detection with SHORE can yield some wrongly detected faces, the false positives (see Fig. 3.4 right). The errors, however, usually relate to quite small areas in the camera picture. In order to filter out the false positives, the framework restricts the size of the recognized face to 2.5% of the whole camera picture. Thus, small room patterns wrongly recognized as a face are eliminated.

Active Users vs. Passive Spectators

Following the approach of Vogel and Balakrishnan [211], the framework divides the space in front of the large display into two zones. The people detected in the nearest zone are treated as active users; they came close enough to demonstrate their williness to interact. People detected further away are treated as passive spectators (see Fig. 3.5). By default the roles are assigned based on proximity. However, if an explicit identification takes place, for instance, by means of a mobile device, the roles are assigned based on mobile identification.

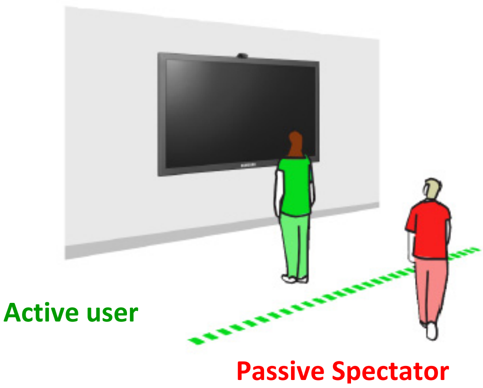


Figure 3.5. Proxemic division of roles: active users and passive spectators.

Proximity to the Display

Proximity context can be retrieved either using a depth sensor, such as Microsoft Kinect [103], or by the analysis of faces detected by a camera.

The depth sensor delivers immediate coordinates of the detected person, mapped on three axes. The data of the z-axis (depth) can be used as an absolute value or mapped to a certain value range, such as “near”, “medium”, “far” distance. The ranges can be used, for example, to define the proximity areas where the person transitions from a passive spectator to an engaged spectator or a user (see Fig. 3.5).

The proximity data can be also retrieved from the faces detected by the camera. The data can be is calculated using the coordinates the outlining rectangulars for each detected face. Supposing that the similarly distant people have similar sized faces, the distance of each face can be derived from the square of the outlining rectangular. If A is the top left coordinate of the rectanglar, and B – the bottom right coordinate, the square equals:

$$\textit{Face square} = (B.x - A.x) \times (B.y - A.y)$$

The larger the square, the closer the person stands to the display.

Since the camera an be fixed at an arbitrary position in the room, the exact translation of the face size into the distance from the display must be done experimentally. In the setting that was used in the lab, the experimentally derived formula estimates the distance between the user and the face in meters as:

$$\textit{Distance} = 120 \,/\, \textit{Face square.}$$

Here the value 120 is an experimentally derived coefficient having no units. Face square is calculated in square pixels, it depends on the resolution of the camera.

Mutual Proximity between Spectators

The calculation of the mutual proximity of spectators is also based on the outlining face rectangular. However, not only the square of each face rectangular is estimated, but also the distance between the faces on x-axis.

Gender

Gender estimation is supported by SHORE software. For each detected face the software estimates the probability of the face being male or female. The values of either probability are attached to the output vectors, together with the face ID and the coordinates of the outlining rectangular.

Age

Age estimation is also delivered by SHORE. For each detected face, the software delivers a value of estimated age and the probability of the estimation. The age of the highest probability is attached to the output vector, together with the probability rate.

Group Constellation

Proximity to the display, mutual proximity of spectators, their gender, and age are combined to derive the group context. If detected people are on a similar distance to the display and they are located close to each other, they can be seen as one group. This assumption holds when the display has a large size, and the space in front of the display is sufficient to stay further away from each other. If no space limitations are present, and people still hold together, it is likely that they belong to one group. Important to emphasize the limitation of our group recognition approach: the group recognition is not sensitive to a group, where individuals stand far from each other. For instance, the constellation of a three distinctly standing persons will be recognized as three single individuals; whereas the constellation of two closer standing persons and one person afar from them will be recognized as a group and a single individual.

The group can be characterized by the age of the group members and their age. Thus, the system can classify the spectators more flexibly. For example, two detected spectators can be two teenage boys, a couple, or a mother with a kid. In many cases the ability of the system to distinguish between such groups is essential for taking a right adaptation decision.

Emotions

Emotions are recognized by SHORE software, for each face detected by the camera. Four kinds of emotions can be recognized: happy, sad, angry, and surprised. Every detected face is analysed for similarity to each of the four emotions, based on the facial expressions. Probability of each emotion is attached to the output vector.

The framework allows to set a threshold for the acceptance of the emotion decisions. Currently, the threshold of 80% is set. If an emotion is estimated by SHORE to be more than 80% probable, the emotion is accepted as true. Otherwise, the emotion of the face is set to neutral.

User Identity

The framework supports the recognition of a user's identity in two ways: explicitly and implicitly. The explicit identification is enabled by mobile devices. The users announce their presence in the ambient room by logging into the system. The advantage of the approach is that the system can be sure about the correctly recognized identity of the user. The disadvantage, however, is that the system cannot trace when the user leaves the system. Thus, if an identified user disappears without logging out, the system will assume that the user is still present.

The implicit identification is based on face recognition. In order to enable the face recognition, the system needs to be trained in advance. The integrated face recognition algorithm is based on the OpenCV software [159]. Once a face is detected in the room, the system tries to match it to the already classified faces. The match with highest probability is attached to the output vector together with the probability rate.

Social Relationships

Social relationships as a context is derived from the user identity context. The framework relies on a pre-defined network of relationships, a graph connecting existing users. The framework distinguishes between friends, friends of friends, and strangers. Friends are users directly connected to each other in the network. Friends of friends are the users connected not directly but via one person. Strangers are the users not connected to each other at all.

Conversation

The microphones deliver a constant audio signal that reflects the noise in the display surrounding. The framework digitalizes the signal by means of nAudio [150] open library. The resulting output volume is presented in a normalized value in the range between 0 and 100. The framework subdivides the volume into three zones: quiet 0..30, medium 31..75, and loud 76..100. The boarders between the zones were chosen experimentally. They can be adjusted depending on the loudness of the public place.

In order to conclude whether people are engaged into a conversation fact, the framework processes the output volume in the following manner. If the volume level raises from the zone quiet to medium or loud, the system starts observing the volume. If the volume level keeps the same level for at least 3 seconds, the system sets the conversation flag as true. Otherwise, if the volume level within 3 seconds sinks to quiet again, the system resets the conversation flag. It starts the observing 3 seconds again when the level raises from quiet to medium or loud.

Such mechanism was developed in order to eliminate the arbitrary noises of a public place: a tram passing by, a noise from the other room, etc. A conversation next to the display usually has a constant volume level, higher than quiet. Moreover, it is unlikely to have breaks longer than 3 seconds.

User Activity

The Ambient Sensing Framework interprets user activity as an activity related to a display. Thus, touch input on a touch surface as well as physical manipulation of a device is registered as an activity.

In the given setting, touch events are registered on the tabletop, the tablet PC, and the mobile device. The framework sets the user activity on device as true once the touch event was performed. Then the timer starts and runs for one minute. If within one minute no touch events are registered, the activity on the device is again set to false.

The physical manipulation of the device is registered using speed and acceleration data. These values are derived from the gyroscope data. Both tablet PC and mobile device have intergrated gyroscopes.

The integrated gyroscopes measure the spatial position of the device related to x-, y-, and z-axes (see Fig. 3.6). The measuring is performed every 0.5 seconds. The gyroscope values are delivered in the range from -1 to 1. Calculating the change of the device position on each axis, one gets the distance the device has passed over time. Dividing the distance by the time, one gets the speed of the device motion on the axis. Dividing the speed by the time, one gets the acceleration with which the device moves along the axis.

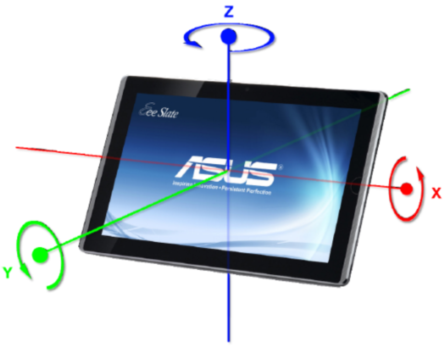


Figure 3.6. Recognition of acceleration data on a tablet device.

Attention to the Display

The attention context is derived differently for every display type.

Public display concludes about spectators attention based on: the presence of spectators, their proximity to the display, the frontal head orientation towards the screen, and their conversation activity at the display. These context data are delivered by camera and microphone.

Tabletop concludes about spectator's attention based on: conversation next to the table, touch events. Conversation context is enabled by the microphone integrated into the table.

Tablet PC and mobile device conclude about spectators attention based on: conversation next to the device, touch events, and physical motion. The data are enabled by the integrated sensors: microphone, touch surface, and gyroscope.

Important to mention, the attention context needs to be supported by time component. The recognized social context must stay valid at least for few seconds, in order to be interpreted as a valid attention. For instance, a frontally oriented head of a spectator will be considered "attention" only if the person stays at the display for some period of time. Taking into account the time component the framework excludes spontaneous observations or conversations of passers-by who are not engaged with the display content.

Further in this work the visual attention will be used as a proof of the spectator's *interest*. The notion of interest in this work equals to a spectator's "visual interest". A person is considered interested in the content, if he/she spends some time observing the content. Such an approach eliminates the confusion with the person's intrinsic interest or commercial interest. Although the question of a person's real interest versus "visual interest" is definitely an important point, it is out of the scope of this work. The methods to measure audience's interest can be found in the related literature [142]. An argument against the "visual interest" can be the necessity of people to stand in front of the display, for instance, while waiting. The frontal orientation of the observing face, however, does imply the visual interest of the observer who keeps his or her face oriented to the content.

3.1.4 A Show Case: Detecting Attention with the Ambient Sensing Framework

In order to illustrate how the framework can be used in a multi-display environment, a demo application was developed. In the demo the framework was used to detect spectators' attention mapped to the devices in a multi-display environment.

Attention context was recognized as described above. Different kinds of displays used different sensors and derived social context in order to conclude about attention.

The system processed the events coming from the sensor and generated the blocks of raw data. The collection was performed on each device involved into the system. The collected data was sent to the main module, where the raw data was translated into the immediate context. Every time a new block of raw data was coming from the devices, the main module compared the previous state with the new state, represented by the data from the blocks. If the previous state ws not equal to the new one, the main module updated the state and sent out the update – the change of the social context.

In order to demonstrate the attention detection by the framework, a trial was conducted in a lab. For the trial, the attention context was represented as the display colour. If a display got the spectator's attention, it was coloured in green. If the display lost the attention, it became blue (see Fig. 3.7).

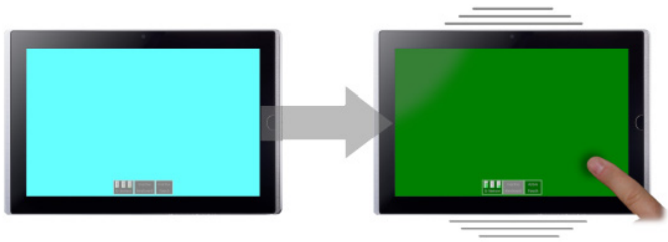


Figure 3.7. Visualization of spectator attention: once a display got attention, it was coloured in green.

The trial was conducted with three test participants. The persons were leading a conversation, moving within the lab, joining each other, and standing alone. The displays – a public display, a tabletop, and a tablet PC – were involved into the scene (see Fig. 3.8).

Figure 3.8. Demonstration of attention detection in a multi-display environment.

The attention detection was running on the involved devices all the time. In order to estimate the correctness of the detection, the subjectively observed attention was compared to the attention fact detected by the system. The subjectively observed attention was derived differently for different display types.

For the tabletop, the real touch events were compared with the detected touch events. The conversation activity was taken into consideration as an additional factor supporting the attention; the touch was considered the main attention criteria. The correctness of recognized touch events reached almost 100%.

For the tablet and mobile devices, the touch events and the spatial motion were considered. Comparison of real touch events with the detected touch events yielded almost 100% of match. The spatial motion, i.e. the presence of the devices in the hands of the users, was also detected almost in all cases. The recognition rate, however, depends on gyroscope sensitivity of each device, which in our case was very fine.

For the public display, the correctness was calculated based on the detected frontal faces. Important to emphasize that the frontal orientation of the face was taken as an assumption of the "visual interest" and therefore was taken as the ground truth.

The camera mounted at the top frame of the public display yielded in recognition accuracy of 90%. Among very few false positives, there was about 10% of false negatives. The false negatives were mostly caused by the slightly tilted down faces, when the spectators stood very close to the screen. Such faces do not appear frontal to the camera mounted on the top pf the large display. In order to fix the recognition errors, the camera must be placed lower. However, the location of the spectator so close to the large screen is rather rare: the comfortable observation distance for large displays starts with 1.5 meters. At such distance the highly mounted camera can catch the faces even with a slight inclination.

3.1.5 Applicability and Extension of Ambient Sensing Framework

The Ambient Sensing Framework aims to provide an application with diverse data on social context. The applications can use the provided social context for different kinds of adaptation.

One of the potential applications is a privacy protecting adaptation. As an example, consider a system that shows some content on the large screen. By means of the framework the system catches changes in the social context: number of spectators, proximity of spectators, and user identity. If the system detects that a new spectator is coming closer to the screen, it performs the necessary adaptation. Additionally, the system can check the identity of the spectator. If it is a friend of the user, the adaptation will be adjusted.

Another potential application is the solution of space conflicts. Imagine several users actively interacting with a tabletop or a public display. The framework notifies the system that a new user is approaching the display. Knowing the mutual proximity of the users, the system can sense from which side the user approaches the display and thus provide a space for the user.

Adaptation of content according to the group interests is another possible adaptation. By means of the framework the display can recognize the group standing in front of it: the number of people, their age, and gender. Knowing the interests of different groups, the display can flexibly change the content entertaining its spectators. These and further examples appear in the following chapters of this work.

The framework is flexible for extensions. Modern sensor technologies develop rapidly, offering new and faster possibilities for social context recognition. One of such extensions can be done with precise motion capturing with Microsoft Kinect [103] or more sophisticated systems such as Sixth Sense [199].

3.2 Sensing the Social Context with Mobile Sensors

The Ambient Sensing Framework enables flexible recognition of context in a room equipped with ambient sensors. A big limitation of the framework is the fixed nature of the sensors: the sensors are mounted within a certain area (e.g. a room), and thus the framework delivers the social context relevant only to the interaction within the area. In order to sense the social context within an entire house, a street, or a city, a huge number of sensors would need to be installed. Such an equipment and complex synchronisation of sensor data is difficult to imagine.

A more elegant solution is to "mount" the sensors directly on the people, the ultimate producers of the social context. For this, every citizen should carry the sensors which would save the data on social context and send it to the central system. The distribution of personal context data is, of course, controlled by the person and is perfomed only when agreed by the user. The only question arising is how to make citizens wear the necessary sensors? How to spread the sensors among the numerous citizens?

The answer indeed already exists: mobile devices are perfectly suited for the role of the sensors; they are widely spread in all corners of our planet. Thus, if the numerous citizens share their social context data with the central system, the system will possess the truthful and updated data on the social situtaion in a house, a city or on the entire planet.

This idea inspired the creation of the Mobile Sensing Framework. The framework collects the social context from individual persons by means of mobile sensors. Below I present the overview of the existing mechansm for real-time data sensing with mobile devices, introduce the Mobile Sensing Framework, and provide an illustration how the framework can be used on urban public displays.

3.2.1 Existing Approaches to Mobile Context Recognition for Public Displays

Despite the advance of sensor technologies available today, the social context collected by the mobile users is not widely exploited [77, 158]. Looking at the research done in this direction, one can notice that urban displays mostly use the Bluetooth sensor.

Memarovic et al. proposed an entertaining content for public displays in the city of Oulu, which relies on the number of Bluetooth devices detected around the display [139]. The content of the display generates quizzes and curious facts using the number of detected Bluetooth devices.

Mahato et al. identified Bluetooth devices in the range of an urban bar in order to retrieve the visitors' music preferences [131]. The music playing in the bar was adjusted to the tastes of the visitors.

Mobile Bluetooth sensors are often used to retrieve the interests of the persons around an advertising or information public display [6, 209, 101]. Sensing the interests of the surrounding people, the display dynamically adjusts the current content or advertisement.

Besides Bluetooth sensor, many other mobile sensors can contribute to the collection of social context [7]. For example, modern mobile devices possess integrated accelerometers, microphones, compasses, GPS, and even NFC sensors. The trend goes further to physiological or bio-sensors. Indeed, bio-sensors become more invisible and aesthetically appealing; perhaps soon people will see tiny heart-rate sensors integrated in users' ear canal [93] or EEG sensors integrated into casual headphones [93].

In spite of a great potential of these sensors to gather social context [171], the sensors alternative to Bluetooth are still rarely employed. The ideas to integrate bio-sensor data into interaction design mostly belong to artists.

An interactive installation “Pulse Room” of Rafael Lozano-Hemmer presents 300 light bulbs which pulsate in the heartbeat rhythm of the exhibition visitors [167]. When a visitor holds the interface of the heart rate sensor, his or her pulse is recognized by the computer; the signal immediately sets off the closest bulb which starts flashing at the exact rhythm of visitor’s pulse.

Heartchamberorchestra is a visual performance accompanying a musical concert [84]. The animation on a large display installed on the stage is controlled by the heartbeat of musicians.

To summarize, the potential of the mobile sensors to collect social context of the people is underexploited. In spite of a wide variety of available wearable sensors, the existing systems mostly take advantage of wireless and Bluetooth sensors.

3.2.2 Mobile Sensing Framework

The Mobile Sensing Framework enables to collect social context by means of various mobile sensors. The framework exploits the wireless, gps, Bluetooth, and gyroscope sensors, a microphone, as well as the external wearable sensors connected to the mobile device.

The framework architecture is based on a server which accumulates and interprets the context, and mobile clients which collect of the sensor-based context.

The server receives the data from each client. The data contains information about the user location, user identifier, surrounding loudness, user motion, and user pulse. An activated mobile client runs in the background; it permanently senses the data of the microphone, accelerometer, and a pulse sensor connected via Bluetooth. In the current implementation, the mobile clients send the data to the server every minute.

Below the author describes in detail how mobile clients collect the context data.

Client ID

The client or user identifier represents a unique identifier of an active mobile device. The number of all activated client identifiers gives information about the amount of people in a certain location. Moreover, if the user sets the user identifier as non-anonymous, the real identity of the user can be discovered. The latter function is utile for the social networks: it enables to find out more information about the current location and state of the one’s friend circle.

Location

The location data reflects the coordinates of the current user location. The value can be sent either in the format of latitude and longitude, or as an identifier of a specific location, such as a name of a bar or a square. The latter format can be used for a special purpose, for example, to trace the social context in selected city locations.

Once a mobile client is activated inside a city location, it establishes wireless connection to the location's WLAN and receives the unique identifier of the place [176]. Given the city locations are tagged with their maximal capacity, the clients identified in the same location give information about the density of the city location.

Loudness

The microphone of the mobile phone registers the loudness of the place. The loudness is given as a numeric value from 0 (absolutely still) to 100 (very loud). The total loudness of a location is calculated as an average of all loudness values incoming from mobile clients in the given location.

The current prototype faces issues with filtering out the noise from the real volume values. The noise is especially significant when the mobile device is placed inside a pocket or a bag, and the person moves. The problem can be solved if the system measures the loudness not by mobile sensors, but with a microphone installed directly inside the location.

Motion

The motion of the user is identified using accelerometer data [172]. The oscillations of a mobile device are observed during the timeframe of one minute. The user's motion is classified into one of the states based on oscillation amplitudes. For example, the states can be "sit", "stand" or "dance". The minimal amplitudes refer to the resting state or sitting. The irregular and small oscillations refer to standing. The larger and repetitive oscillations refer to dancing. The threshold values of oscillation amplitudes were set experimentally.

The identified motion patterns ("dance", "stand" or "sit") are sent to the server. Thus, the overall motion pattern in a location is estimated by counting the number of clients with motion values "stand", "sit", or "dance".

In fact, the recognition of human motion activity using accelerometer data is usually done with machine learning algorithms [115, 172]. Our algorithm based solely on amplitude thresholds can be improved in this way. Machine learning is necessary to distinguish complex motion patterns such as walking, running, standing up or sitting down. We, however, have only three motion categories which can be sufficiently precisely classified using the oscillation patterns. Of course, this approach should be improved if more patterns are to be recognized, for example, distinguish dancing from walking or running.

Pulse

The pulse of the user is transmitted by a mobile heart rate sensor. The current version of the client exploits a remote pulse sensor, the wearable Zephyr HXM sensor [220]. The sensor sends the pulse data to the mobile client via Bluetooth. The sensor represents a light belt-bound sensor which does not require any user input. The only effort required from the user is to wear the sensor belt. The future mobile devices are expected to have an integrated pulse sensor, thus eliminating the necessity to wear a separate sensor.

The sensor transmits the pulse data to the mobile device every second. The mobile client calculates the average pulse for the last minute and compares the value with a prerecorded range of the visitor's resting pulse. If the current average exceeds the upper border of the resting pulse, the mobile client reports the pulse as "higher" than average.

Alcohol Consumption

The data on alcohol consumption can be derived from the pulse data. The pulse data can be provided either in an absolute value, or indicate the trend of the user's pulse: it can have either the value "average" or "higher". If the pulse is detected as "higher", the visitor is either moving actively (dances) or consumes alcohol. If the pulse is detected as "higher" than average and the motion is detected as "still" or "standing", the alcohol consumption is identified as true.

Figure 3.9 summarizes the social context recognized by the Mobile Sensing Framework.

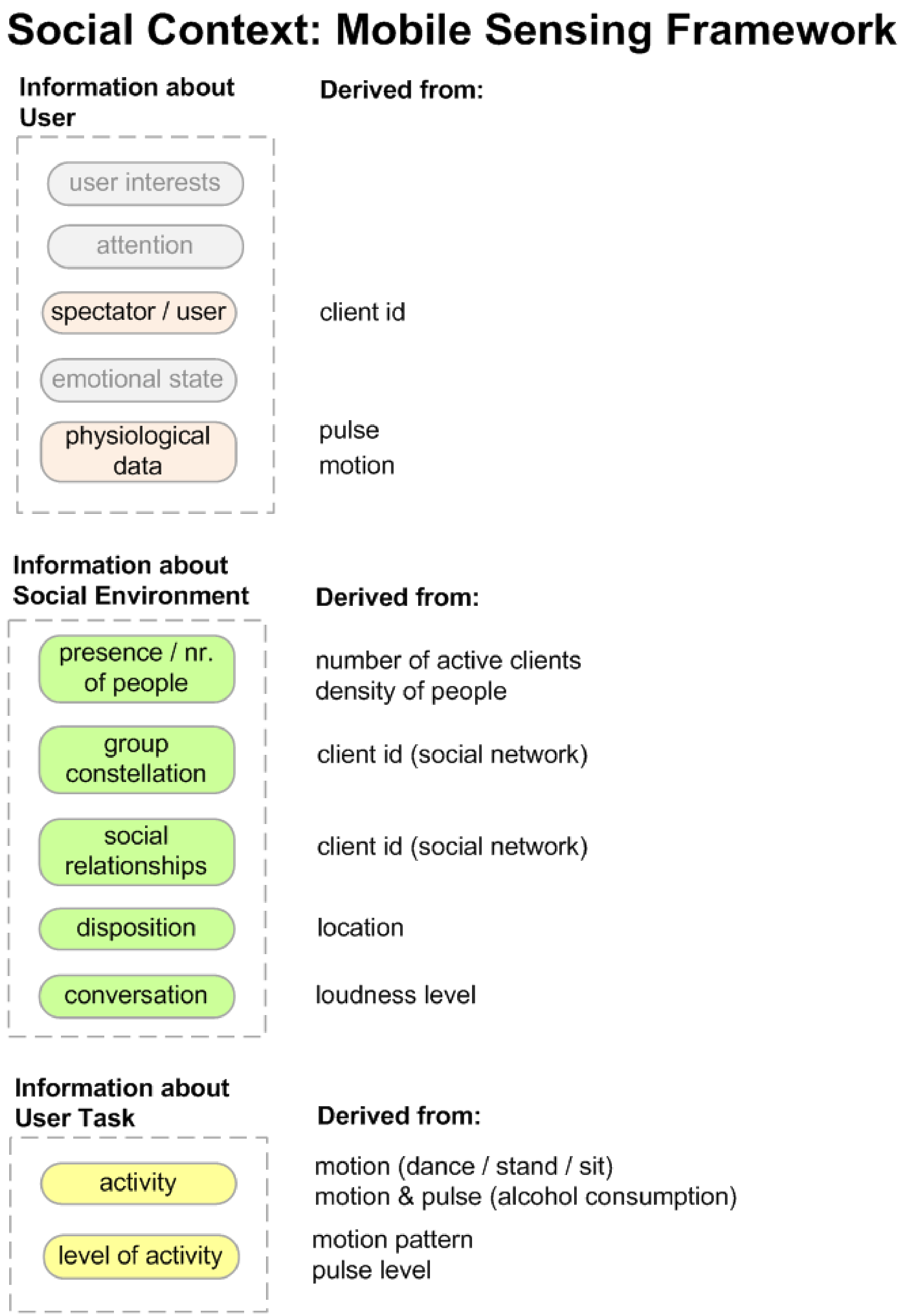


Figure 3.9. Social Context recognized by the Mobile Sensing Framework.

3.2.3 A Show Case: City Pulse Urban Display

In order to illustrate the functionality of the Mobile Sensing Framework, the "City Pulse" prototype has been created. The prototype represents a public display which reflects the immediate social activity in city locations (bars, cafes, restaurants). It helps passers-by plan or adjust their going out activities while being already in the city. The large display installed in a central square of the city renders information collected by the mobile sensors: the crowdness of the locations, the loudness inside, whether people dance or not, and how much alcohol is consumed.

The display utilizes the concept of Web 2.0 [160]: the real-time context data collected from the visitors of the city locations. The advent of Web 2.0 [160] has significantly improved the way people plan their leisure time activities. Relying on the user-generated content, such as Flickr [60] pictures or online forums, people eased and diversified the way they choose the locations for travelling, sightseeing or going out. However, relying only on the online content, people often risk finding the information outdated or incomplete. For example, a bar in the top-ten list is temporally closed for renovation or does not exist anymore. The City Pulse public display aims to solve this problem. The mobile context brought on the public display enables a quick overview of the situation in the city right now: the spectators can estimate how crowded or loud the locations are, whether people dance or not, whether they drink or not.

Presented approach offers certain advantages. First, the data is collected in an implicit way. Unlike existing technologies, such as foursquare [62] or Facebook [55], the mobile clients of the visitors do not require any explicit contribution, such as a manual input of their location. Second, the data can be gathered in an anonymous way. The display gathers solely quantitative information, without any link to the personality of the content contributor. Finally, the public display is placed directly in the city, where people mostly need an immediate support in finding a going-out location.

Apart from the reflection of the current situation in the city locations, City Pulse can highlight the locations matching personal preferences of the user. For instance, it can show the bars with jazz music or restaurants with Japanese cuisine. Moreover, the display can highlight the locations where the user's friends are currently going out.

3.2.4 Designing City Pulse

When choosing a going out location, people usually follow such criteria as the density of people inside, loudness, whether people inside mostly dance, stand or sit, whether people inside are drinking, price category, distance from home, the style of the ambient music [154].

In order to visualize this data, the author reviewed the existing approaches for the visualization of quantitative data [207]. The visualizations for urban analysis often use density maps [105, 57], or heat maps [67, 68, 66, 63], emphasizing the key information with shape, size, and color of the highlighted spot. Bergstrom et al. encode critical information in the saturation of the spot [19]. In order to show dynamic information, displays can use animations. For instance, the ambient installation of Rafael Lazaro presents the light bulbs animated with the rhythm of visitors' heart beat [167].

The resulting design of City Pulse is illustrated on Figure 3.10. The display provides the main view the map of the city (or city part) overlaid with the social context information. The downtown map in shown schematically in black and white; the context data is shown in colour. The context data obtained from the citizens is depicted by coloured spots overlaid over the map.The social context includes the data on the density of people, loudness, people motion (dancing, standing or sitting), and alcohol consumption.

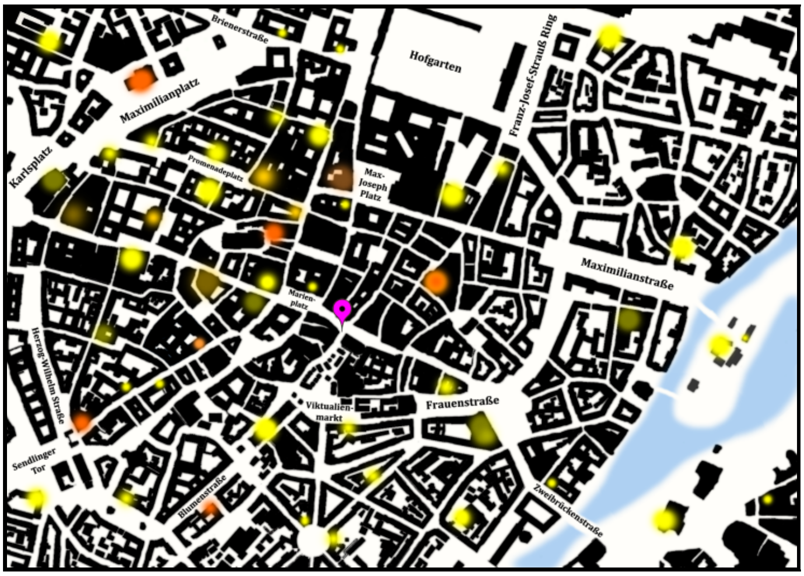


Figure 3.10. City Pulse display: downtown of Munich.

Density of people inside the location is mapped to the size of the coloured spot: the smaller the spot, the less density inside. If the spot goes out of the building silhouette, the place is overcrowded.

Loudness is mapped to the opacity of the spot. The opacity of 100% corresponds to a very loud place where it is difficult to speak. The opacity around 0% makes the spot almost invisible. Indeed, if the loudness is registered as zero, the place is absolutely empty and is probably closed. Thus, there is no need to display it on the map. Crowded but quiet places, such as restaurants, can be distinguished on the map large semi-transparent spots (see Fig. 3.11, left).

Motion of people is shown through an animation. If at least 30% of the visitors are dancing, the City Pulse map animates the location spot with flashing pulsating, so "making the spot dance". The 30% ratio – or a third of visitors – is enough to consider the location with dancing people inside. Thus, the display can signalize to the passers-by who are in the mood to dance that they will not be the only ones who dances. If the ratio of dancing people is less than a third, the system estimates the proportion of standing and sitting visitors. The places where the visitors mostly sit correspond to static spots. If the majority is standing, the spot is pulsating slowly.

The *alcohol consumption* in the given location is calculated by counting the number of visitors with pulse “higher” than average and motion not equal to “dance”. The consumption of alcohol is encoded in colour. A place with low alcohol consumption is depicted with yellow spot. The redder the spot, the higher alcohol consumption (see Fig. 3.11, right).

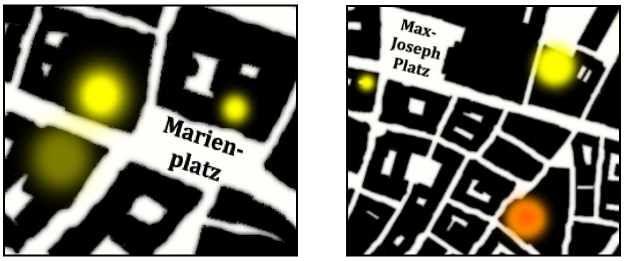


Figure 3.11. Loudness is depicted by opacity of the spot (left), alcohol consumption is encoded in colour (right).

The users of the City Pulse display can make personalized requests, such as finding a location according to their taste, or finding their friends. City Pulse display highlights the silhouettes of the buildings which correspond to the matching locations (see Fig. 3.12).

Figure 3.12. Highlighted locations matching a personalized request.

The request to highlight the friends is processed using a social network. This functionality so far is implemented with a prototype social network, but can be exchanged for a real one using for example Facebook API. In order to locate the user’s friends, the system finds the clients related to the client identifier of the requesting user. If such friend’s client is activated in the city, the dots are rendered over the corresponding locations.

A request to highlight the locations of friends yields the display of coloured dot. Each dot stands for one friend, hiding however any information on the friend’s personality (see Fig. 3.13).

Figure 3.13. Highlighting friends in city locations.

The personalized requests are enabled by a mobile client. The personalized requests can be done by multiple users simultaneously. In this case the users will be assigned different colours of icons, for example, magenta and green.

The visualization on the display was implemented in C#, using WPF platform. The mobile clients were implemented on Android 2.3 platform. Each client had a unique client identifier, preset during the installation.

3.2.6 Applicability and Further Extensions of Mobile Sensing Framework

Mobile Sensing Framework collects the mobile sensor data to reflect the social situation in the city. The context data however can be used for any other purpose. For instance, it can be used for counting people on public gatherings, such as demonstrations or festivals. It can be also useful to trace people's movement within an area, such as a city or a region. The main advantage of the system over the existing approaches is the access to immediate context, the context which reflects the situation now.

A future extension of the approach can be seen in the integration of a recommender system. The display may initiatively propose the users the locations based on the user's previous going-out history, user's preferences or preferences of user's friends. The recommender system can also consider the environmental context, such as the weather, events in the city or the daytime. Thus, in a cold winter day the display would rather propose cozy cafés, and in a hot summer evening – an opened terrace or a beer garden. Finally, the recommender system may take into account the emotional and social context. For example, it would propose a loud social place with many friends inside, when the user is sad and alone; or a romantic place when the user is happy, accompanied by a girlfriend.

3.3 Summary

This chapter presented two approaches for the recognition of social context. The first approach enables context recognition with ambient sensors installed in a room. The displays located in the room can thus provide detailed information about the surrounding spectators.

The second approach collects the social context by means of mobile sensors. The mobile devices of citizens collect data using integrated microphones, wireless sensors, accelerometers, and heart rate sensors. Given that the entire city carries such devices, the system can provide comprehensive real time information about the social context in the whole city.

Comparing two presented approaches, one can highlight the advantages and drawbacks. The mobile data collection is more flexible, it does not limit the area of context recognition and thus can be scaled for cities, countries, and continents. However, the approach works only under the condition that a sufficient number of the people possesses recognition-enabled mobile devices.

The ambient data collection provides more possibilities for recognition. Apart from camera and microphone, other sensors can be quickly integrated into the room, for instance, motion sensors. People inside the room do not have to carry any additional equipment in order to enable the recognition. Neither do they have to make any explicit input to start the recognition. However, the context recognition is running only inside the room where the sensors have been installed, thus leaving the social context outside unknown.

Chapter 4

Designing Adaptation driven by Group Context

Personalized content on public displays offers clear advantages: the users get direct access to the information of their interest. A challenging task, however, is to *learn the interests* of the users and to *offer matching content* in real time. These tasks become even more complex if the displays are installed in busy public places, where numerous individuals circulate every day. Such displays have to learn the interests of a huge amount of users; moreover, they need to combine the interests of distinct individuals, when a group of several individuals observes the content.

Another challenge relates to the implicit nature of the personalization mechanism: the content adaptation should happen automatically, without any input from the user side. Indeed, a manual input, such as activation of an online profile or switching on a mobile client, is hardly acceptable in a busy public place. People have no time and attention for the manual input; moreover, they may be just unaware of the input possibility.

Therefore, there is a need of a mechanism that would study the interests of the spectators and adapt the content completely *automatically*. Such a mechanism should distinguish between *different profiles* of spectators, however, requiring *no explicit input* from the spectators' side.

The system proposed in this chapter meets the described requirements. In order to illustrate the concept, imagine the following scenario. A busy train station is equipped with a large public display. The display provides incoming passengers with city-related advertisement and useful tips. When a mother with two kids passes by the display, the screen advertises leisure activities for families. When a single lady passes by, the display shows a trendy perfume shop. For a passing-by couple, the display shows tips on romantic cafés. Finally, for a group of teenage boys, the display advertises an adventure attractions park.

The learning of spectators' interests and the real-time adaptation are enabled by the recognition of *group context*. By means of a camera mounted on the display, the system scans the composition of the observing groups: the number of individuals, their mutual disposition, and gender. Additionally, the system registers positive emotions of the individuals and whether they have a conversation.

In the learning phase the system tracks visual attention of the groups and relates it to the popularity of the content categories. During the real-time adaptation, the system recognizes the group standing in front of the display and shows the content which has the highest popularity within the given group.

By means of an experiment, I demonstrate how the system can be used to identify differences in the visual attention (interest) of different groups [112]. Important to mention, the notion of interest in this work equals to a spectator's "visual interest" (see Section 3.1.3).

After an overview of related research, the personalization mechanism is described in detail. In order to prove the usefulness of the system, the author presents the experimental deployment conducted in a public area of a university. Although this public space is definitely less busy than, for example, a train station, it does represent a valid public space with active and irregular circulation of diverse individuals. Therefore, this space is suitable to prove the performance of the system. The results of the experimental deployment show that the system can be successfully used to tag the content according to the group-based observation patterns (visual interest). Moreover, it provides interesting insights into the diversity of circulating groups and their emotional and conversational activity while observing the presented content.

4.1 Personalization of Content

In order to provide personalized content in real-time, a system should (1) learn the interests of the spectators and (2) be able to present the right content in real time according to the learnt interests. The first task is usually achieved by *tagging*. The second task refers to *real-time adaptation*.

4.1.1 Tagging

By means of tagging, the content elements are labeled with the spectators' interests. The interests can be retrieved in an explicit or implicit way.

For *explicit tagging*, a sample of potential users is asked to rank the presented content manually [97]. Such a tagging is usually done in laboratory conditions. The users rate each element of the content according to a given schema, e.g. a linear scale. Although the method delivers exact results, these results might not reflect the *real* user interests. Since the lab setting is not natural, the ranking may deviate from the preferences the users would express in a real setting. Moreover, explicit ranking requires a significant effort from the user. The tiredness caused by the ranking routine, therefore, may also impact the number of provided ratings and the reliability of the results [97, 168].

Explicit tagging is barely applicable in a public place scenario. Since a great number of individuals are circulating in a public place, it is unrealistic to suppose that each of them contributes to the tagging of the displayed content. The ranking rather will reflect the interests only of some distinct individuals who did take their time to contribute. Therefore, the ranking can significantly deviate from the interests of the numerous other individuals circulating in the crowd. Another disadvantage refers to the extraction of group interests. Complex algorithms must be applied to derive the group interests from the interests of distinct individuals [134]. The complexity grows in busy public places where the group compositions are usually very diverse.

Implicit tagging is a more adequate ranking approach for public places. The method usually exploits crowd monitoring. An illustration of implicit tagging can be found in the work by Müller and colleagues [144]. The method counts the number of frontal faces registered for every content element (e.g. a slide). The element which has accumulated the largest number of the faces is considered to be of the highest interest. In the subsequent real-time adaptation process, the "most interesting" element will be set to the highest priority in the content schedule.

The approach replicates the real behaviour of the spectators: people look at the content when they are interested in it. However, the approach is not flexible enough to distinguish between groups of spectators. What if one content element was observed only by numerous single persons, and never – by couples? The approach, however, will prioritize the content element for both groups: singles and couples.

Another disadvantage of the method is the assumption that the frontal look equals interest. Although visual attention is an important hint to derive interest, it is not sufficient. In fact, the spectators can look at the display for many other reasons [90, 143]. For instance, the display is oriented frontally to the spectators' path or the colours of the content subconsciously attract attention. Such effects do not necessarily imply interest. Therefore, for automatic personalization more contextual cues should be used to support the assumption of interest.

To summarize, implicit tagging is the most suitable approach for public spaces. Such tagging requires no user input and reflects the natural setting. The tagging mechanism must distinguish the compositions of various spectator groups.

4.1.2 Real-time Adaptation

The second part of the personalization process refers to the real-time content adaptation. The adaptation is based on the results obtained by the tagging. The real-time adaptation can rely on user *contributions* or work completely *automatically*.

The *contributions*-based approach requires a certain registration from the user side. For instance, the spectators switch on their Bluetooth devices and transmit the pre-set profiles [6, 101, 131]. The display receives the profiles and adjusts the content according to the profile interests. By means of dedicated strategies [96] the group interests can be derived from the individual profiles [209]. Although the contributions-based approach provides a precise overview about the present spectators, it is hardly applicable in a crowded public place. The numerous visitors of the public place may not possess the required devices. And even those who possess them may simply forget to switch the device on. As a result, the display will retrieve an incomplete or a wrong picture about the surrounding spectators.

The *automatic* approach can utilize the identity of the user. For example, by means of face recognition the system can understand who stays at the display, and thus automatically adjust the content [211]. Although such an approach does not require any user contribution, it has to carefully learn the user profiles in advance. This requirement is not realistic in a busy public place with numerous individuals. Müller and colleagues [144] used face detection as a trigger for real-time adaptation. Once a face is detected in front of the display, the most popular content appears on the screen. The method is more suitable for public places than identification: it eliminates the unrealistic knowledge of each single individual. However, it does not distinguish between the interests of various individuals. For instance, the content popular among women does not necessarily match the content popular among men. Moreover, the method cannot does not take into account the composition of the spectator groups. For example, the same content will be displayed for two teenage girls, a couple, a group of elderly men or a mother with three kids.

All in all, it is a challenging task to recognize the interests of spectators and provide the right content in the real-time. Ideally, the system should know the individual profiles of the spectators, which is difficult to realize in a busy public place. Moreover, the system should work automatically, not requiring any input from the spectator side. A trade-off would be an approach which approximates to the individual profiles, but does not require the spectators to explicitly provide the system with their profiles.

This work presents a group-based personalization approach, focusing on the tagging phase. Based on the gender and disposition of detected spectators, the system classifies the spectators into distinct groups, for instance, a couple, two men or a single woman. Thus, the system registers the interest not of distinct individuals, but of distinct groups. Since the detection of the spectators and the recognition of their gender are done with a camera, the approach works fully automatically eliminating any user input.

4.2 Approach for Group-based Personalization

Based on the analysis of the existing approaches for tagging and real-time adaptation, the author came up with a set of requirements for personalization systems in busy public places:

- the system must work **completely automatically**, not expecting any contribution from the user (spectator) side.

- the system must **capture the group composition**, instead of identities of distinct individuals. The unique groups should be defined by the characteristic features of the group members, e.g. gender or age.
- the system must be **applicable in a busy public place**, taking into account the diversity of passing-by individuals and possible groups.

The system presented in this work meets the requirements of automatic personalization. Below the author describes how it can be used for automatic tagging and real-time adaptation.

Completely Automatic System

Content tagging and subsequent adaptation are implemented using the Ambient Sensing Framework (see Section 3.1). The context used in the current work is summarized in the Figure 4.1.

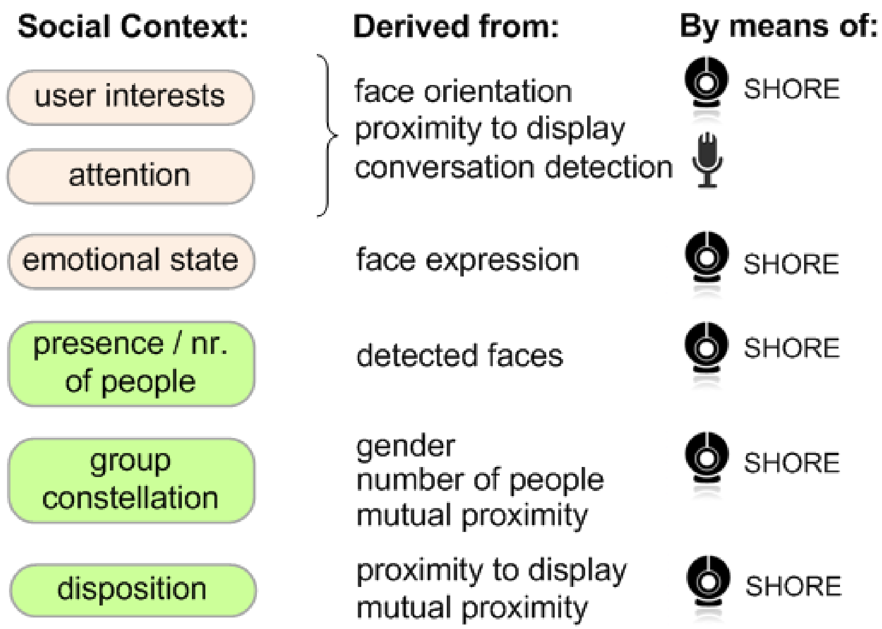


Figure 4.1. The social context recognized by means of the Ambient Sensing Framework.

The framework was utilized to recognize the group of the spectators currently observing the display content (thus, having "visual attention"). Each group was characterized by the number of individuals, their gender, emotional state, and the conversational activity. The volume level is classified into three ranges: silent (almost no sound), moderate (moderate discussion between several individuals), and loud (active discussion). The proximity information was used to identify the group constellation: whether the present people belong to the same group or there are some alone standing individuals. The details on the recognition of either context data can be found in the Section 3.1.3.

The Figure 4.2 illustrates the output of the camera picture: the detected faces are marked with coloured rectangles (red for female, blue for male) and the key points of the face features.

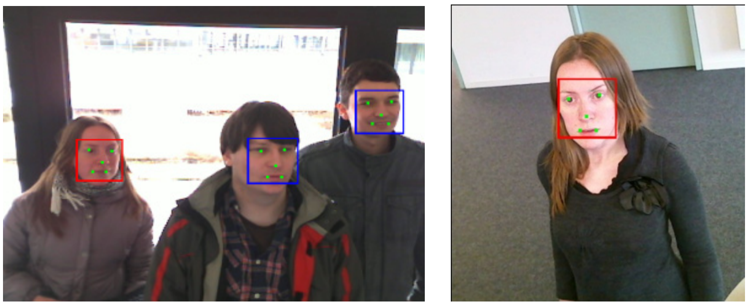


Figure 4.2. Recognition of the group structure: red and blue rectangular indicate female or male faces.

Capturing the Group Composition

Having the data on present faces, their gender, and disposition, the system can conclude about the group composition. The spatial disposition of the faces enables us to determine whether the present spectators belong to the same group or are standing alone individuals.

Applicable in a Busy Public Place

Generally, the number of SHORE-detected faces is not limited. The software, however, allows defining a minimal size of the face outline as a percent of the entire camera field. The author set the smallest face to be 2,5% of the field covered by the camera. Smaller faces refer to distant persons who cannot see the content properly; thus, they are not considered as valid spectators.

The recognition of spectator's interest can be supported by additional cues, such as positive emotions or discussion of the content. Our personalization mechanism registers emotions of each group member, as well as the volume level of the conversation. However, the reliability of these additional cues has to be proven experimentally. Emotions and conversations are not necessarily caused by the display content. Therefore, the author considers emotional and conversational response as a secondary hint to the spectator's interest.

Speaking about emotional response, it is important to mention that interest does not always imply positive emotions. For instance, a person can be highly concentrated on the content (thus, interested), but have a neutral facial expression. A positive emotion therefore is not equivalent to "relevant", but is a contextual condition that influences how "relevant" the content is.

4.3 Accuracy of Recognition

Before the deployment of the system, the author tested the accuracy of face detection and emotion recognition delivered by the SHORE. Important to emphasize, the test goal was not to verify the accuracy of the SHORE algorithms. This question has been elaborated by the authors of SHORE and can be found in the related literature [114]. The test goal was solely to test how accurate the SHORE software performs in the given experimental conditions.

The test was conducted in the public area of the university, where the main experiment took place, employing the same displays as in the main experiment. For the test the author presented some arbitrary photos on the displays. All spectators were video-recorded; a note informed them about the recording fact. Simultaneously with the video recording, the tagging system was running in the background. It logged the detected faces, their gender, and emotions.

In total, 16 hours of video were recorded, containing 128 female and 120 male faces. The video material was manually annotated, registering the recorded number of female and male faces, group constellation, and emotions (based on subjectively estimated facial expressions). The annotation was compared with the log data, yielding the accuracy of recognition.

About 95% of all the faces of people standing more than 1.5 seconds in front of the display were captured by SHORE. The faces further than the specified observation distance (face rectangle covering less than 2.5% of the camera view) as well as the faces of passers-by who just glanced at the display without stopping were not detected by SHORE as faces. This limitation is in line with our definition of spectators: people within a close proximity to the screen, who do stop to watch the content. Recognition precision of the face detection approaches 90%; the detailed data on precision are given in the Section 3.1.4.

Gender Recognition of SHORE showed accuracy rate of 96% for males, and 92% for females. The system needed about 0.3 seconds on average to decide on the gender.

Emotion Recognition. All in all, the system recognized 2557 “neutral” emotions and 2526 emotions other than “neutral”. Out of 2526 non-neutral emotions, there were 1983 “happy” emotions, 100 “angry” emotions, 186 “surprised”, and 257 “sad” emotions. The recognized emotions were compared with the faces recorded in the video. As a result, the accuracy of recognition was calculated. The recognition accuracy of “happy” emotion yielded 90% for male spectators and 92% for female spectators. The recognition accuracy of “surprised”, “sad”, and “angry” emotions showed less reliable results, yielding only 60%, 61%, and 65% respectively. These emotions were recognized on the video rarely: usually, spectators either smiled or didn’t express any emotions.

To summarize, gender recognition with SHORE can be reliably used for the group-based personalization. As for the emotion recognition with SHORE, one can reliably apply only the recognition of the classes “happy” vs. “not happy” emotion (in other words, “smile” vs. “no-smile”).

4.4 Experimental Deployment

The goal of the experimental deployment was to see how well the system can be applied for group-based personalization. In particular, the following questions were to be answered:

- Can the system identify differences in observation patterns (content preferences) of distinct spectator groups?
- What insights into spectator groups can be gained with the system?
- Does information on emotional and conversational response deliver reliable hints on spectator interests?

4.4.1 Deployment Set-Up

The system was deployed during three weeks on the displays in a university public area. Three displays were involved in the deployment: one display situated in a lobby and two in a passage (see Fig. 4.3). All displays have non-touchable screens of 62 inches and 45 inches in diagonal.

Figure 4.3. The lobby display (left) and the passage displays (right) used in the experiment.

The circulation of people on the premises of the university is moderate. Besides the main "inhabitants", consisting of about 30 researchers, the experiment area is used by students and visitors. The passage area is often used as a short cut to the university canteen, the parking lot or other places within the university. During the experimental weeks two events took place at the area adjacent to the experiment public place; bringing in total about 30 visitors from outside the university.

The aim of the personalization system was to tag the content newly created for the university displays. The content was compiled in a slide show; the personalization system ran on the background. The content topics were proposed by researchers of the university. Within a brainstorming session, the researchers came up with four content categories: "Team", presenting the members of the research team, "News", informing about recent info, e.g. upcoming events or lectures, "Department Life", presenting events of research unit, "Quiz", posting a tricky question about a research unit, followed by the correct answer. The researchers found these categories relevant for the university life. However, the author needed to find out whether the content would also attract our students and visitors.

The design of the content was kept consistent, in order to exclude distractions caused by visual design (see Fig. 4.4). Each content slide stayed on the screen for 10 seconds.

4.4.2 Tagging Procedure

In order to tag the content categories according to the group interests, the tagging algorithm was launched in the background of the slide show. Each display was supplied by a camera with an integrated microphone. The cameras were installed on the top side of the display frame. The SHORE software was processing the images captured by the camera. With the frequency of 15 frames per second, SHORE delivered information on each detected face: gender estimation, coordinates of outlining rectangular, and emotion estimation.

Figure 4.4. Examples of the content: “Department Life” and “Quiz”.

This data was processed to make an entry to the log file. Based on the number of detected faces and the gender data, the <**group composition**> was calculated. Based on the coordinates of the rectangular, the author calculated <**position**> of each group member. Position reflected the user location at the display (left, centre, right) and the proximity to the display (near, middle, far). From this information the author could estimate whether the spectators belong to the same group (stand next to each other) or are several distinct individuals. Important to notice, the proximity information can be used only as a hint for the group division. Comparison of the recognized groups with the real groups in the recorded videos yielded quite high recognition accuracy 92%. However, the result can be explained by the low circulation of spectators. In the most cases the spectators appeared at the displays either alone or with a friend. Different groups almost never crossed at the display. Therefore, the proximity-based group devision is applicable for the areas with rather low circulation of people. The exact numbers and the diversity of the groups are given below in the Section 4.5.

Based on the probability of each emotion, the author registered the resulting <**emotion**>. The emotion having probability more than 80% was entered to the log. Finally, the microphone provided data on estimated <**volume level**>.

As a result, a log entry consisted of the general description of the social context and the detailed description of each face:

<timestamp> <group composition> <volume level>

<face1><gender><position><emotion> <face2><gender><position><emotion>

An entry was added to the log every time the social context was changing, for instance, people joined the group, people left, or emotional context changed. The following lines illustrate an example of a log entry (F stands for female, M stands for male):

<15:29:00> <2F + 1M> <loud>

<face1><F><left near><neutral>

<face2><F><left near><neutral>

<face3><M><left near><happy>

The log files were created for every day, separately for each display. The tagging system did not capture any raw video and audio signals.

After the experiment, the log files were parsed. The author summarized how frequently the groups observed each content category, which emotions were expressed, and the volume of the conversations. Additional information, such as the number of all groups, the total number of females, etc. could also be derived from the log files.

4.5 Experiment Results

In total, 324.2 hours and 4727 detected faces were recorded in the log files. The analysis of the log files enabled to answer all questions posted to the experiment. First, the experiment proved that the system is able to recognize the interests of distinct spectator groups. Second, the author obtained interesting insights for the groups circulating in the public area. Finally, the author could conclude whether emotional and conversational context can support the evidence of spectators' interests. Below the author provides the detailed results.

The system successfully identified the differences in observation patterns (visual interests) among distinct groups. Figure 4.5 illustrates the distribution of visual interests among groups of two or three spectators. The illustration clearly shows the differences in observation patterns: for instance, topic "News" was more frequently observed by the group "1 Male + 1 Female" than by other groups. Topic "Quiz" was more often observed by homogeneous groups, "2 Males" or "2 Females". The distribution refers to the data obtained at the lobby display; very similar distribution patterns were observed on the passage displays. The figure reflects the interests of only composite groups; the interests of single spectators (one male or one female) were distributed similarly to the interests of the respective groups of two (two males or two females).

In total, the system detected 10 different kinds of groups. The majority of the detected spectators were single individuals (see Fig. 4.6). This finding was quite surprising for us, since many meetings and collaborations take place at the university. Observing the behaviour of spectators, the author realized that in spite of the gathering in meeting rooms or lecture halls, the transitional public places (such as the passage and the lobby) people mostly pass alone. Detected composite groups consisted mostly of two persons.

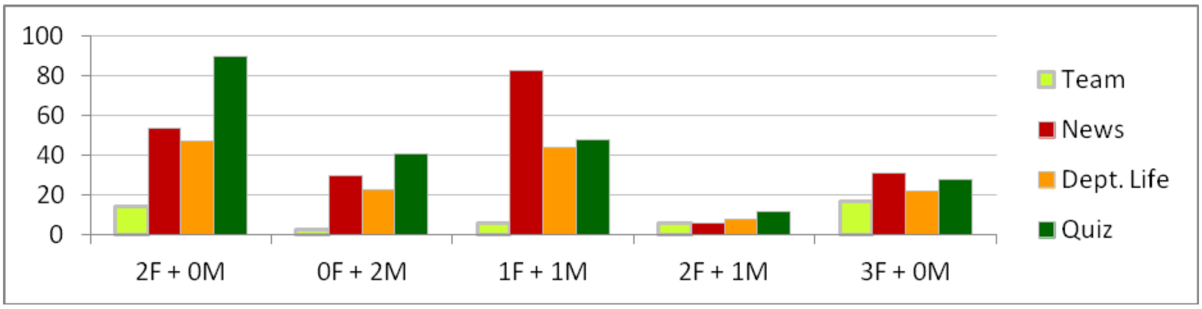


Figure 4.5. Distribution of group interests (F stands for female, M – for male). Y axis indicates the number of times the visual interest of the group was detected.

During the experiment the system registered a solid number of positive emotions. For the analysis, only positive emotions were considered, since the SHORE software yields reliable recognition results only for "happy" vs. "not happy" emotions (see Section 4.4).

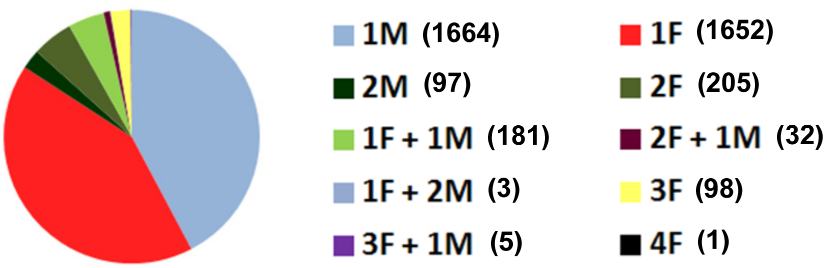


Figure 4.6. Distribution of spectator groups. M stands for male, F - for female.

Figure 4.7 illustrates the distribution of positive emotions among single female and single male spectators. The analysis embraces single males and females, since they were the most represented spectator types. The analysis for other groups can be done similarly.

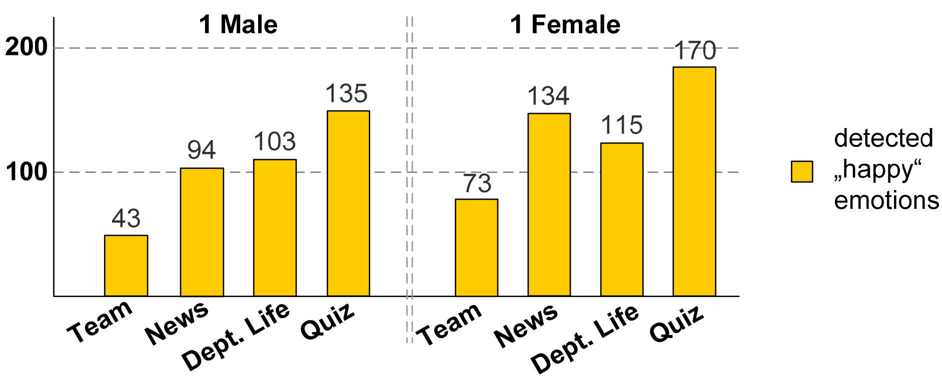


Figure 4.7. Distribution of "happy" emotions.

From the first sight, Figure 4.7 uncovers clear differences in the frequency of positive emotions expressed for different content topics. However, calculating conditional probabilities for each topic (considering how frequently each topic was observed by either group) the author didn't find any noticeable differences.

Male spectators showed almost equal emotional response to all content topics: "Team" (0.34), "News" (0.34), "Department Life" (0.36), and "Quiz" (0.33). Females had a lightly more frequent positive response to "Team" (0.53) and "News" (0.45); however, quite a similar response to "Department Life" (0.4) and "Quiz" (0.39). Generally, males expressed positive emotions slightly less frequently than females (0.34 and 0.43).

The analysis of the conversational activity was done in a similar way. For the analysis the author considered the groups of two spectators. The author chose these groups for the analysis, since the system mostly detected conversations between two persons. For each group the author calculated conditional probabilities: how often and in which volume a group had conversations while observing the content.

The analysis did not reveal any noticeable differences. Most of the conversations were done in moderate volume, independently on the content topic.

Homogeneous groups (only males or only females) were slightly more silent when observing the content "Quiz". Mixed groups had generally slightly more conversations when observing "News". These observations can be explained by the nature of the content. "Quiz" posts the spectator a question, substituting a real conversation and thus making people silent. "News" provokes a discussion about some urgent events. However, the conversations could also be not related to the content.

4.6 Discussion

Below the author provides the interpretation of the experiment results, addressing the research questions posted above. Moreover, the author discusses limitations of the study, further steps, and possible applications of the presented approach.

4.6.1 Content Preferences of Distinct Spectator Groups

The experiment has shown that the group-based personalization mechanism can successfully extract the differences in observation patterns of distinct spectator groups.

The main interest differences can be observed between homogeneous groups (only males or only females) and mixed groups (a male and a female). Homogeneous groups mostly preferred "Quiz" category, whereas the mixed groups were more interested in "News".

The phenomenon can be explained by the relationships within homogeneous and mixed groups. Observing our spectators, the author noticed that homogeneous groups often represent close friends. They meet at the university not only for study-related occasions, but also for socializing, chatting or spending a free time slot together. Therefore, they are likely to involve into such an entertaining occasion as a quiz. Mixed groups often represent study fellows, connected not by a friendship, but rather by a common studying activity. They meet at the university for a certain study-related occasion, e.g. to work on a project or prepare for an exam. Therefore, they are not likely to spend time for a “Quiz”, but would rather pay attention to the study-related “News”.

The overall majority of spectators showed more interest to the content “Quiz” and “News”. The preference to “News” relates to its informative content: people tried not to miss relevant and important facts. The preferences to “Quiz” can be explained by its interactive nature. The quiz questions were related to the university stories. Therefore, the quizzes not only challenged the spectators, but also gave them some curious facts.

Comparing these results with the observations of Rist and colleagues [98], who evaluated various contents at university displays, the author may see slight contradictions. The authors reported that people *generally* have lower interest in entertainment content and higher interest in news. However, in their work, entertainment related to games which uncover user participation. Unlike games, our entertaining quiz allows the users to participate unnoticeably, with no demonstration of success or failure. Such unnoticeable interaction is known to be appreciated by people in public locations [152, 85].

4.6.2 Gaining Insight into Spectator Groups

Detected groups contained slightly more females than males. The fact was found surprising: statistically our technical institute counts more males than females. One explanation of this phenomenon can be the natural curiosity of women and their ability to notice the surrounding objects better than men [164].

Analyzing the log files, the author could see that single spectators were often joined by other persons, creating a group. Such behavior is known as the “honey pot effect” [135, 148]. People are not courageous enough to demonstrate their interest in public. Thus, they feel more comfortable to join an existing spectator.

Among the 10 detected groups, only 5 groups were presented in the passage area. The circulation of people in the lobby is indeed higher, since it is a large recreation room where people usually gather. The passage, on the contrary, is a narrow corridor. People usually pass it quickly, heading to a certain room or to the canteen. The lower number of spectators in the passage can also be explained by the orientation of the displays. As mentioned by Müller et al. [149], the displays oriented at 180 degrees to the user trajectory attract less attention than the displays oriented at 90 degrees. This observation applies to the orientation of the experimental displays: the passage displays are oriented at 180 degrees, and the lobby display – at 90 degrees to a typical passer-by trajectory.

4.6.3 Emotional and Conversational Response

Observing arbitrary spectators, the author noticed that positive emotions and conversations are often not related to the content. They are usually brought from a dialogue preceding the display observation. Therefore, the experiment results do not give enough evidence that detected positive emotions and conversations were *provoked* by the content.

4.6.4 Application in Other Public Spaces

Although the experiment was conducted in a public space with rather moderate circulation of people, it demonstrates that the system can be deployed in other public spaces. Apart from the university public space scenario, the system can be applied in an environment with a brighter diversity of groups.

The system installed at a large *shopping mall* can recognize the interests of different customer groups. Unlike existing ambient technologies facilitating shopping experience [140], our system is able to learn the interests of the customers. Based on the learnt shopping interests, the system can advertise the matching content immediately when customers approach the display. In a similar way, the system can be deployed at a *travel agency*. It will help to recognize trends in vacation destinations among couples, single travelers, families, etc.

The system can give an insight into the tastes of the people. Imagine the system installed at a *picture gallery* or a photo exhibition. Tracking how visitors observe the art pieces, the system can conclude which authors and which genres are popular among different visitor groups. Such information could facilitate planning of the future exhibitions.

Finally, the system has a potential to impact the tastes of the people. Imagine the system to be installed at a university "*Open Doors" day*. The "Open Doors" day is an annual event organized by universities, aimed to orient school students in the choice of their future education. A current problem of engineering faculties is a low ratio of female students. The problem is partially caused by gender stereotypes, but partially by insufficient awareness of school girls about the engineering career. A display recognizing social context could increase their awareness. Once girls are recognized in front of the screen, the display can switch to the Engineering content.

4.7 Real-time Adaptation

The real-time adaptation is based on the results obtained in the tagging phase and the mechanism enabled by Ambient Sensing Framework.

The results of the tagging phase are summed up in the matrix of adaptation rules. The matrix consists of the list of the recognized groups and the content topics (see Fig. 4.7). The content topics for each group are ordered by their popularity.

	Priority 1	Priority 2	Priority 3	Priority 4	Priority 5
1F	Topic 3	Topic 1	Topic 4	Topic 5	Topic 2
1M	Topic 2	Topic 1	Topic 5	Topic 4	Topic 3
2F	Topic 3	Topic 4	Topic 1	Topic 5	Topic 2
1F + 1M	Topic 5	Topic 1	Topic 2	Topic 4	Topic 3

Fig 4.7. The matrix of adaptation rules.

The real-time adaptation system is working in the following manner (see Fig. 4.8):

1) The content presentation is running in a predefined idle loop. The content topics appear one after another, changing in a preset time period. The Ambient Sensing Framework is running on the background; it permanently analyzes the surrounding social context.

2) Once the Ambient Sensing Framework detects a face in the range of the display, it changes the context “presence” and sends notification to the adaptation system.

3) The framework additionally analyzes the distance of the detected person from the display. The distance context is also sent to the adaptation system. The distance influences the speed of the adaptation.

4) Moreover, the framework analyzes the group context of the detected people. The group context is sent to the adaptation system. If the group context cannot be recognized by the framework, the empty group context is sent.

5) Once the adaptation system receives the notification about the changing “presence” context, it starts the adaptation. First, it looks up the content preferences for the received group.

6) If the group context is successfully retrieved, the system looks for a match in the matrix of adaptation rules. If the group context was not retrieved, the system keeps its idle content presentation.

7) If the successfully retrieved group is found in the matrix of adaptation rules, the system creates the schedule of the content topics, based on the priority list. The topics will be presented to the spectator group as long as the group stays in front of the display.

8) The system switches the content, according to the prepared schedule. The switch speed depends on the distance of the face from the displays. If the distance is short, the content topic is switched quicker. The exact time of the switch has to be found experimentally for each display; it depends on the position of the display and the architecture of the room (the trajectory of spectators).

9) If the group is not found in the matrix of adaptation rules, the system does not perform any adaptation; it keeps the idle presentation order.

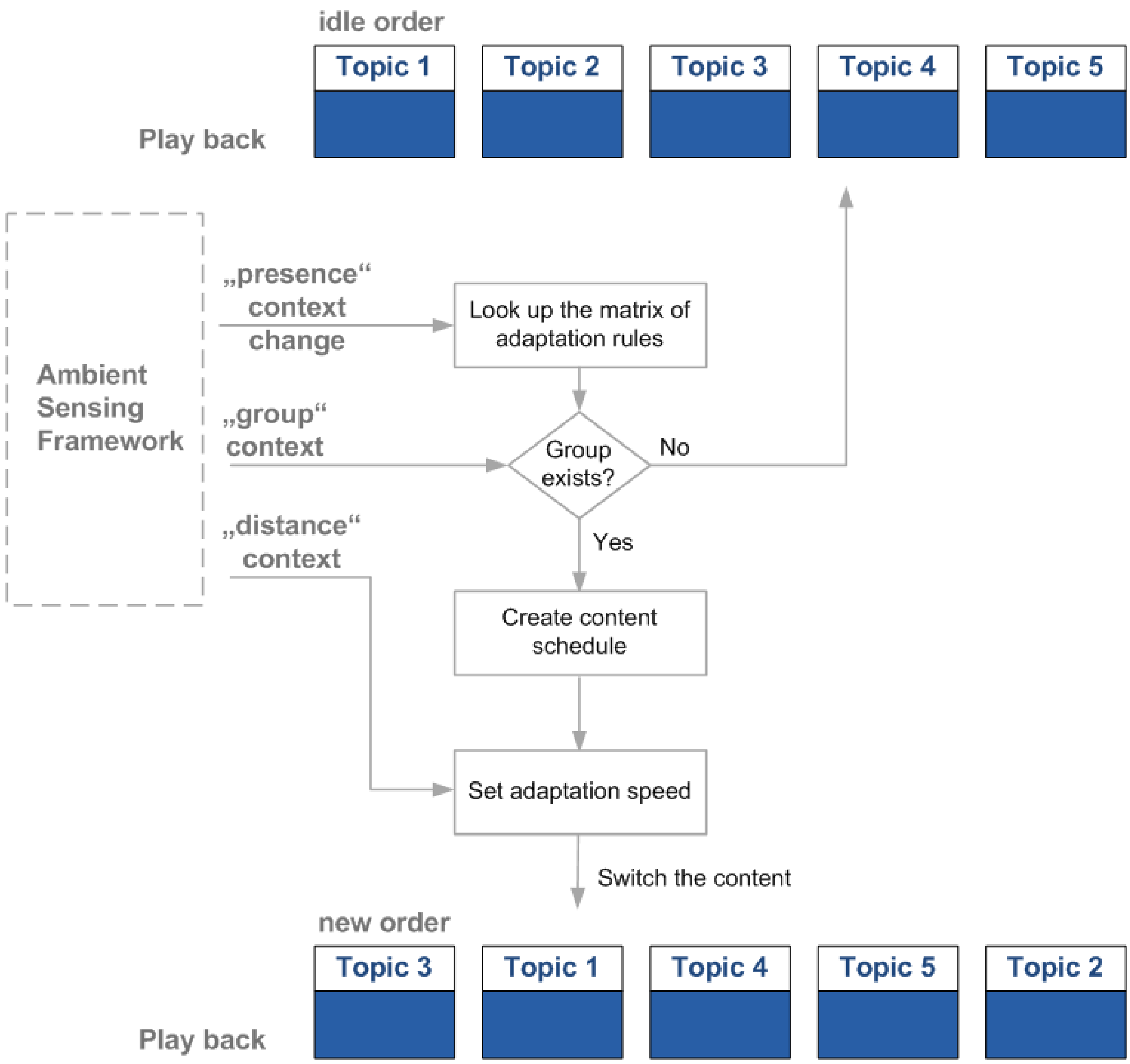


Fig. 4.8. The mechanism of real-time adaptation.

In order to validate the system performance for real-time adaptation, a more realistic *public* setting is necessary. The experiment presented in this work does show that the group-based approach can be successfully applied in a public setting. However, the experiment revealed that the groups presented at the university environment are not that diverse. A real busy public place, such as a train station or a shopping mall, would be more appropriate to test the system in real-time adaptation mode.

4.8 Summary

The chapter presented a system for group-based personalization on public displays. The system can be used for the *tagging of content* according to spectators' interests and for the real-time *content adaptation*.

The advantage of the proposed invention over existing systems is its completely *automatic* adaptation mechanism. The extraction of interests, as well as the real-time adaptation is performed automatically, without any input from spectator side. This requirement is critical in busy public places: the passers-by are unlikely to have time, attention or means for an input.

Another advantage of the proposed system is its capability to distinguish between spectator profiles. Instead of the retrieval of individual profiles (which is hardly realistic in a busy public place), the system extracts *groups profiles*. The groups are defined according to the number of spectators, their disposition, and gender. For example, the display distinguishes between two women, a couple, a girl or a group of boys.

An experimental deployment, conducted in a real public space, proved that the system can successfully identify the differences in observation patterns (visual interests) of different spectator groups. Moreover, the experimental results gave us insights into the circulating groups: what constellations of spectators are typical, which groups circulate in different public areas, what is the proportion of female and male spectators.

Finally, the experiment enabled to conclude whether the tagged data on the emotional and conversational response can be correlated with the displayed content. The results gave us no evidence that positive emotions and conversations are directly related to the content. Often they are caused by events preceding the display observation. However, the author has shown that the system can reliably tag positive emotions. The author believes that a more entertainment-oriented content (such as a photo exhibition) can reveal differences in emotional response. The content chosen for the experiment was rather emotionally neutral; it addressed the topics relevant to the public area – a university environment.

The possible extension of the approach can be done by introducing the *age* context. This advance will enable to distinguish, for example, between two teenage boys, an old couple or a mother with a kid. These groups are likely to have very different content preferences.

Chapter 5

Tailoring Interaction with Public Displays

Interaction with personalized data on a large public display represents a sensitive scenario: first, users expose the fact of interaction in public; second, personalized data may be seen as slightly privacy-critical. In this chapter the author investigates how interaction design can support the user in such a sensitive scenario. Through experimentation, the author compares interaction techniques, various multi-display settings, and user scenarios to find out how social context impacts user preferences in interaction techniques.

5.1 Comparing Direct, Bodily, and Mobile Techniques for Interaction with Public Displays

Interaction with personalized content on public displays brings certain advantages, but also presents risks. On the one hand, users get quicker access to the necessary personal information. The display can automatically tailor its content according to the user profile, helping the user eliminate manual extraction of necessary data. On the other hand, interaction with personal data in public can result in privacy issues. Therefore, the interaction with a personalized display must be designed not only in a usable and comfortable way, but it must also be perceived as trustworthy, reliable, and secure.

The following work presents an experimental study to examine three interaction techniques that are generally used on public displays: direct, bodily, and mobile-based.

Although advantages and drawbacks of the three techniques were widely discussed in related literature [184, 32, 26, 9], the techniques have never been compared in interaction scenarios with a personalized display. A personalized display, like any other public display, exposes its content in a large and comfortable format. The content, however, contains data that reflects on the user's profile. For instance, the display can recommend goods potentially interesting to the user. Alternatively, it can present personal data of the user, which is potentially interesting for a group of observers, for example, a calendar of the user overlaid over the group calendar or pictures from a corporative event.

The work aims at identifying aspects which are critical for fluent and trustworthy (secure) interaction with a personalized public display. In particular, it investigates which of the techniques users perceive as understandable, controllable, comfortable, reliable, privacy protective, and trustworthy. Comparing the results with the insights gained from literature, the author identifies critical design aspects that are specific for interaction with personalized displays. The work is summarized as a set of design recommendations that aim to inform practitioners in designing interaction for personalized public displays.

Although the investigation was done with a personalized display in a lab environment, it gives a clear understanding of the user's perception and preferences in interaction techniques. The other studies comparing interaction techniques in lab environments [184, 32, 26] show that the achieved results represent a useful input for the designers of real-life interactions. The goal of our study, therefore, is to provide an initial input for the designers of the personalized displays in a real world environment. Derived recommendations are aimed to guide the initial stages of the design process when designers are analyzing possible situational context and look for the optimal interaction solutions appropriate for the given contexts.

Despite the sensitivity of a public scenario, the use of personalized displays has become increasingly common [128]. The current study uses two different types of personalized content: The first is an application for visualizing a personal *social network*. The second is an application that presents a *persuasive display*. The user's personal data is encoded in the displayed visualization with to the purpose of influencing user behavior. The examples of similar types of public display content can be found in previous research [44, 91, 126, 85, 179], as well as in real-life projects, for example, the installation of the Interactive Video Wall in Copenhagen [94] or CityWall in Helsinki [165]. The examples show that despite the awareness of privacy issues [128, 85, 118, 89] that can result from sharing the personalized content, people do place their private data on public displays.

5.1.1 Interaction with Personalized Content

Generally, the process of interaction with personalized content on a public display can be subdivided into three phases: (1) identification, (2) navigation, and (3) collecting results.

The *identification phase*, or log-in, is when the user transmits to the display their unique identifier. Once the identifier is recognized, the user's personalized information appears on the large public screen. Identification phase is required in order to interact with the information which is personalized.

The *navigation phase* allows the user to manipulate the personalized data displayed on the screen. The navigation is usually governed by a specific goal. For instance, the user looks for a meeting in their personalized schedule [89], with a desire to save it to a mobile device. The method of manipulation may vary greatly depending on the application: users can browse through displayed items, edit content (e.g. draw), type in a text request (e.g. requesting a train timetable), or even do complex collaborative work. In this work, however, the author limits the scope of navigation activity to browsing through the displayed items. Such manipulations are often employed on outdoor displays for wide-public usage. For example, the Interactive Video Wall [94], in Copenhagen, enables simple browsing through the sightseeing pages. The CityWall [165], in Helsinki, provides browsing through pictures taken by citizens. Another reason to choose browsing, as a navigational activity, is the consistency of the interaction process; the browsing action does not depend much on the particular design solution. The user navigates through the items displayed on the screen, with the ability to highlight and select each item.

The *collecting results phase* refers to the accomplishment of the navigation goal. The goal followed during the navigation phase is usually directly related to the browsing process: the user looks for a specific item to retrieve the necessary information. For example, the users of the Interactive Video Wall [94] look for interesting city sights to get information on the opening hours, shortest route, etc. The users save the retrieved information on a mobile device, such as a personal mobile phone, or send it via email.

Logging-off (removing personal data from the public screen) can be seen as an additional phase. However, since the study focuses on the interaction particulars of the phases, it considers log-off process to be similar to the identification phase. Indeed, for consistency reasons, the log-off is usually designed identical to the log-in (or identification).

It is important to notice that personalized display applications do not always involve all three phases. For instance, the CityWall [165] presents pictures of citizens; users may browse through the pictures (navigation), upload their own pictures (identification), or download existing pictures (collecting results). In this case, each phase is independent and can be skipped. In this work, however, the author aims to investigate user preferences in the three phases. Therefore, our experiment was designed and conducted with applications involving all three phases.

5.1.2 Interaction Techniques on Public Displays

In general, three techniques are usually employed when interacting with public displays: direct, bodily, and mobile-based interaction.

Direct interaction assumes physical proximity between the user's interactive tool and the display. The interactive tool can be the user's hand or an assisting device, such as an NFC-enabled mobile phone. The technique exploits a real-life metaphor of "touching", where an individual activates an object by touching it. The display may enable direct interaction by means of touch-surfaces or by other technologies, such as a matrix of NFC tags [196, 31] or using camera-based hand recognition [194, 76]. Studies on direct interaction show that users perceive this technique as natural [184], fast, reliable, enjoyable, and easy [32]. However, since the interaction is possible only at a short distance, the users have to make an additional physical effort to move closer to the object [184].

Bodily interaction is enabled by spatial gestures, body postures or proximity. The technique is particularly beneficial if the display is located far away from the user and thus it is physically impossible to reach the display. Bodily interaction is usually supported by camera-based recognition [211]. Previous studies show that this technique is quick and intuitive [184]; it increases the user's engagement, enjoyment, and is considered to be fun [127, 21]. However, the user may see gestural interaction as artificial and hard to memorize. As a consequence, gestural interaction can negatively impact the user's cognitive load [108]. In addition, users expressed concern about performing gestures in public: they feel uncomfortable attracting attention of bystanders [85, 152].

Mobile-based interaction utilizes a mobile interface to control the public display. The mobile interface may offer specially tailored tools for control or it can replicate the entire content of the public display [24]. The main advantage of the mobile-based interaction is the ability to interact from any distance, with minimal physical effort. However, users may find the technique boring [32], too technical [184], and inconvenient [184], due to the constant focus switch between large and mobile displays.

5.1.3 Comparing Interaction Techniques

Although each of the interaction techniques described in the previous section were discussed widely in literature, the techniques have never been compared in the scenario of interaction with a personalized public display.

Rukzio et al. [184] compared the techniques touching (direct), pointing (bodily), and scanning (mobile-based) in interaction scenarios with real-life objects. Although the study was conducted in a controlled lab environment, it gives the insights into user perception of interaction preferences. Similar lab studies were conducted by Broll et al. [32] who compared direct and mobile-based techniques in interaction scenarios with paper posters and by Boring et al. [26] and Ballagas et al. [9] who compared mobile-based and gestural (bodily) interaction for controlling a cursor at a distant public display.

Although these works provide interesting insights into the speed, reliability, physical effort, and error rates of the techniques, the results can inform the design of the personalized displays only partially. The sensitive scenario of interaction with personal data applies specific restrictions to the user's acceptance and preferences of an interaction technique. Therefore, there is a need to investigate how the distinct techniques can fit into a scenario of interaction with personalized content. Since the public setting introduces an additional sensitivity condition, to the interaction process, the author needs to find the design aspects that are critical for the use of public displays with this particular condition.

5.1.4 Experiment

To investigate the users' preferences amongst the three interaction techniques (direct, bodily, and mobile-based) with personal data on public displays, the author setup a study. The aim of the study was to identify specific aspects that are critical for the scenario of interaction with personalised display. Moreover, the author aimed to identify how the user preferences in the personalized display scenario differ from the preferences in other scenarios, such as interaction with physical objects or non-personalized public displays.

The experimental study explored the user perception of three interaction techniques (direct, bodily, and mobile-based) used in the various interaction phases (identification, navigation, and collecting data).

Public Display Application

The experiment was conducted with two sample applications, Friend Finder and Late-o-Meter.

Friend Finder visualizes the social network of a user, rendered over a local map (see Fig.5.1). The friends are depicted by icons containing their pictures and names.

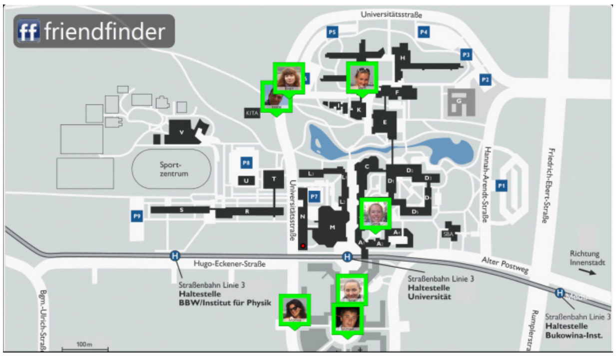


Figure 5.1. Visualization of a user's social network using Friend Finder

The application was designed by students in the framework of a term project [107]. A survey conducted amongst university students revealed that the students wished support in locating their peers on the campus, for instance, to gather for lunch or to work on a project. So far such appointments were done by phone calls or short messages, which was generally found inconvenient and expensive. A display that employs the Friend Finder application and located in a public area, on campus, could help the students to quickly locate their friends.

Friend Finder takes advantage of the large size of the public display. First, the large screen estate gives a good overview of broad social networks. The media such as a map is difficult to observe on a smaller screen, such as a mobile display or a desktop. Second, several users can observe the social network at the same time. This case is widely spread in the student environment: a group of friends may share a friend circle and would like to view the location of their friends together. Moreover, several independent users may merge their friend networks on the same map. In this case, each network is presented in a unique colour of frames. In the study, however, the author focuses on a single-user scenario.

Once the user has loaded the social network on the screen, he or she can browse through the friends' icons, and retrieve the shortest path to the selected friend. The path can be downloaded to their mobile device for later navigation. This function was especially appreciated by the students, since many of them have orientation difficulties around the campus.

Late-o-Meter displays the weekly delays of a group (such as a group of students attending the same lecture course) and aims to persuade the group members to be more punctual. The idle view (see Fig. 5.2) shows anonymous black silhouettes. Each silhouette represents one group member; the silhouette's height is mapped to the person's delay in minutes. Since the identity of the silhouettes is hidden, a passer-by can only see how punctual the whole group is. A group member can personalize the display and see his or her own delays and compare them with the group. The personalized view highlights the user's silhouette in orange and it shows their exact delay (see Fig. 5.3). The user is then able to go into the details of their weekly performance that shows how well they managed their delays to lectures.

Figure 5.2. Visualization of a group delays using Late-o-Meter (idle view).

Late-o-Meter originated from a students' term project. A survey conducted amongst university peers revealed that many students wish to improve their punctuality. In spite of the shame in front of the lecturer and the class, their habit of being late does not change by itself. With the help of Late-o-Meter students believed they can become more punctual.

Late-o-Meter benefits from the large screen: the ambient display provides an overview of the group success and helps students fight for punctuality as one team. Personalization enables students to see their own success and compare it with the rest of the group. Personalizing the Late-o-Meter, users can see the details only of their own delays. The rest of the group remains anonymous. Thus, the users can compare their delays with the delays of the others respecting privacy of the group members.

In spite of the slightly private character of the personalized data, the students appreciated both applications and expressed their willingness to use them. This positive feedback is in line with the observations of other research projects dealing with personalized displays: despite awareness of privacy issues [128, 85, 118], caused by the personalized content, people do use public displays to interact with personalized data [128, 94, 165].

Interaction Phases

The interaction phases (Identification, Navigation, and Collect Results), for each application, are described in the Table 5.1.

	Friend Finder	Late-o-Meter
Identification	User brings the personal social network on the large screen.	User's silhouette is highlighted in orange.
Navigation	User browses through the friends, selecting the friends' icons on the map.	User views the daily performances, represented by bar graphs, selecting the exact delays per day.
Collect Results	User saves the shortest path to the currently selected friend.	User saves the memo of the selected day.

Table 5.1. The three interaction phases in Friend Finder and Late-o-Meter.

Design of Interaction Techniques

The Friend Finder and Late-o-Meter applications were developed in three versions that individually support one of the three interaction techniques (direct, bodily and mobile-based).

Direct technique presumed contactless interaction with visual markers that are displayed on the large screen. The contactless interaction was enabled by a mobile phone. In an initial version of the application, the direct interaction was supported by camera-based marker scanning with an Android mobile device. The user had to capture a marker integrated into the content of the large screen. Once the mobile camera has recognized the marker, it sent the command to the server, and thus the necessary action (e.g. identification or selection) was triggered on the large screen. Although the camera-based scanning was reliably working, the marker capturing procedure caused unacceptable interaction delays. Therefore, for the experiment the author used the Wizard-of-Oz approach, imitating an immediate contactless touch.

The users interacted with the visual tags in all three phases. For identification, the user made a "contactless touch": bringing their mobile phone close to the log-in marker that is displayed at the bottom of the screen (see Fig. 3).

Figure 5.3. Identification with direct interaction.

The user identification succeeded through the recognition of the unique ID stored on the mobile device. Once the user was identified, the personalized data (social network or the personal delay) appeared on the large screen.

To navigate through the items, the users selected the items by the “contactless touch” directly at the item location. In the Friend Finder application, the users had to touch the icon of the friend, while in the Late-o-Meter application, they touched the marker associated with the representation of a day graphical bar (see Fig. 5.4). The collecting result and log-out were performed similarly to the identification: the user had to touch a marker at the bottom of the screen.

One can argue that the described direct interaction is indeed a mix of direct and mobile techniques, since the user touches the tags with a mobile phone. However, the mobile technique in our understanding implies an active interaction with the mobile screen. The presented direct technique on the contrary exploits the mobile phone only as a tool to enable the touch. Such tool can be substituted by the user’s hand, finger or a pen [179].

The current implementation is though more complex than a hand interaction, however, it can be applied to a wider range of the displays which are not equipped with a touch surface. The interaction by physical touch supported by marker-scanning is more universal and realistic; it can be provided by any display assisted by any camera-equipped mobile phone.

Figure 5.4. Navigation, direct interaction: Friend Finder (left) and Late-o-Meter (right).

Bodily interaction was enabled by the Ambient Sensing Framework. Figure 5.5 summarizes the context used for the experiment.

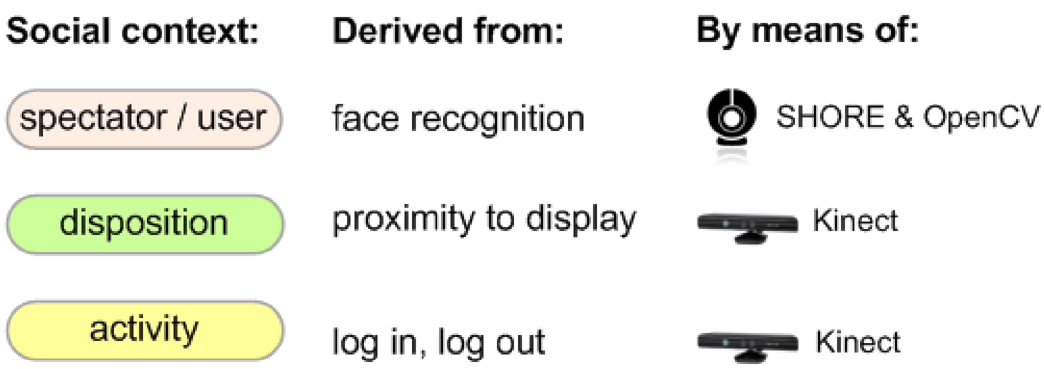


Figure 5.5. Social context provided by the Ambient Sensing Framework.

For identification, face recognition was used together with the proximity context [76]. In order to log-in, the user had to come closer to the display and cross a certain proximity border (1.5 meters) (see Fig. 5.6). The user identity was recognized by means of face recognition. To log-off, the user had to step back behind the proximity border.

Figure 5.6. Identification with bodily interaction.

For the experiment, the author disabled the individual face recognition, in order to avoid extensive training of all test participants. For simplicity reasons the identification succeeded once the face of a participant was detected to be inside the proximity zone.

Navigation through the content was supported by MS Kinect, where the user's hand coordinates are tracked using depth sensors. In order to select an object, the user had to point with the right hand at the respective item on the large screen (see Fig. 5.7).

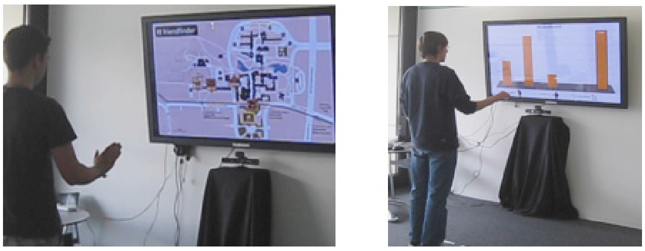


Figure 5.7. Navigation with bodily interaction.

Collecting results was also supported by MS Kinect: in order to save the path to a friend (Friend Finder) or save a delay memo (Late-o-Meter) the user had to raise their left hand.

Mobile-based interaction was supported by a mobile client running on an assisting mobile device. The client for Friend Finder was implemented on a Windows Mobile device; the client for Late-o-Meter was implemented on an Android phone. In both applications, the users could log-in, log-out, and collect the result using a respective button. The identification succeeded once the unique mobile ID was sent to the display.

Navigation, however, was designed differently for the Friend Finder and the Late-o-Meter applications. Generally, it is a challenging task to design a "blind" mobile phone control for interaction with a large screen; so that the mobile interface minimizes the uncomfortable focus switch between two heterogeneous displays.

In the Late-o-Meter application, such a blind control was easier to design. The linearly arranged day delays on the large screen can be spatially mapped to the similarly arranged mobile buttons (see Fig. 5.8). The mapping is made by pressing a button on the mobile screen, which in return activates the respective bar on the large screen.

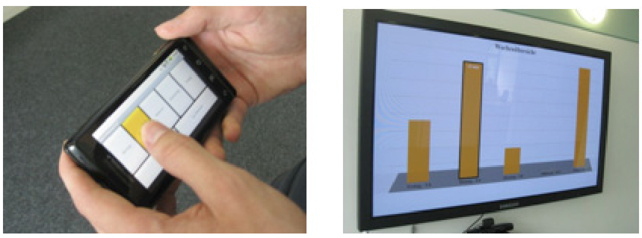


Figure 5.8. Mobile-based navigation in Late-o-Meter.

In the Friend Finder application, locations of the user's friends on the map represent rather an unordered structure. The icons of the friends are placed according to the different locations of their friends, which can also change over time. Following user-centered approach, the author evaluated several controls for navigation through the icons [107]. The final version of the circle-based control was inspired by the iPod-wheel [8] (see Fig. 5.9).

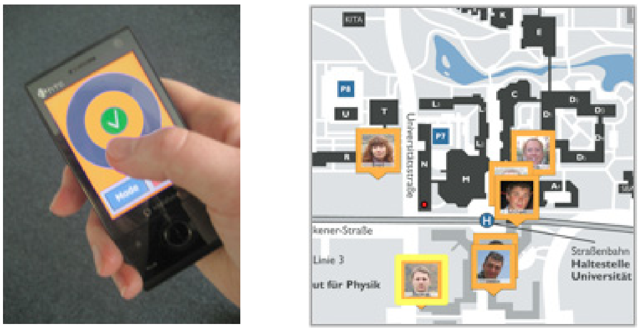


Figure 5.9. Mobile-based navigation in Friend Finder.

By looking at the disposition of the friends on a large display, the user can arrange the friends' icons into an imaginary circle. Scrolling the mobile wheel in either direction allows the user to select friends, one by one, located at the current navigational angle on the map. This technique was seen more convenient and quick than tabulation or navigation with arrows.

Experimental Design

The experimental study was aimed at finding how users perceive different interaction techniques in both applications (Friend Finder, Late-o-Meter[1]). It aimed to identify the similarities, understand the differences, and explain the user preferences for each of the interaction phases (identification, navigation, collecting data).

The experiment was conducted as a between-groups test, in order to exclude learning effects caused by experiences with other application. One group evaluated only Friend Finder, the other group evaluated only Late-o-Meter.

Within either group, participants were evaluating all three interaction techniques: direct, bodily, and mobile-based. The order of the techniques was counterbalanced.

The test was conducted individually, in a public area of a university. After a short introduction about the experiment and the applications, every participant was given a task: they had to log-in, select several items one after another (friends for FF and days for LoM), collect the result (path to selected friends and delays for a selected day), and to log-off. In every task, the routine was repeated three times, to assure that the participants got sufficient interaction experience. The same social network and delays were used for all the participants.

Each participant had to go through the task using the three versions of the application: direct, bodily, and mobile-based. After conducting the task for each version, participants were asked to fill in three questionnaires: for identification, navigation, and collecting results phase. Each questionnaire aimed to capture how well the current interaction technique supported the given phase.

[1] Here and further, FF stands for Friend Finder, LoM stands for Late-o-Meter.

Based on the work of [184, 32, 26, 152], which compares different interaction techniques, the author derived questions that focus on investigating six design properties: how the users find the interaction technique in terms of transparency, controllability, comfort of use, reliability, privacy protection, and trust.

The questions were formulated as statements that the participants had to rate on a 5-Likert scale, from "strongly agree" to "strongly disagree".

Q1: *The system behaviour was comprehensible*

Q2: *I had control over the system behaviour*

Q3: *It was burdensome to use the system*

Q4: *I found the system reliable*

Q5: *The system appropriately protected my privacy*

Q6: *I found the system trustworthy*

At the end of the experiment, the participants were asked which technique they would prefer in each phase and why.

Participants

In the experiment, a total of 34 students participated in the study (17 Friend Finder, 17 Late-o-Meter). There were 10 females and 24 males, aged from 21 to 36 (mean 28.7), engaged in IT, Law, and Literature. All participants had previous experiences with mobile devices; 18 of them were experienced in bodily interaction (mostly from entertainment games), just 7 participants had experiences with contactless technologies. None of the participants was familiar with FF and LoM.

5.1.5 Experiment Results

This section provides the results obtained from the statistical analysis of the questionnaires. The results per application were analyzed using paired samples t-test.

Although the experiment was conducted with two different personalized applications, user preferences results for the interaction techniques in Friend Finder and Late-o-Meter were surprisingly very similar. The following sections describe the results.

Identification Phase

In the identification phase, 56% of participants gave their preferences to mobile-based interaction, 32% preferred bodily technique, and 12% chose direct technique as a preference (see Fig. 5.10).

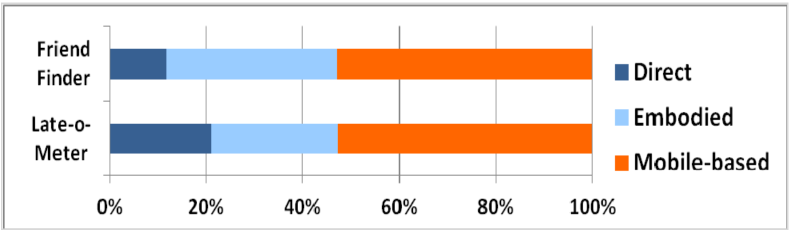


Figure 5.10. Distribution of preferences: Identification phase.

Mobile-based interaction was perceived more controllable than bodily interaction (FF: t = -3.1, df = 16, p = 0.0069; LoM: t = 3.85, df = 16, p = 0.0014). Mobile-based technique was also found more privacy protective than the direct technique (FF: t = -1.77, df = 16, p = 0.095, LoM: t = -1.95, df = 16, p = 0.069). Mobile-based interaction was also found more reliable (FF: t = -1.9, df = 16, p = 0.028) and more trustworthy (FF: t = -2.4, df = 16, p = 0.029) than bodily interaction.

Navigation Phase

In the navigation phase, 65% of the preferences were given to direct technique, 26% to bodily technique, and only 9% to mobile-based technique (see Fig. 5.11).

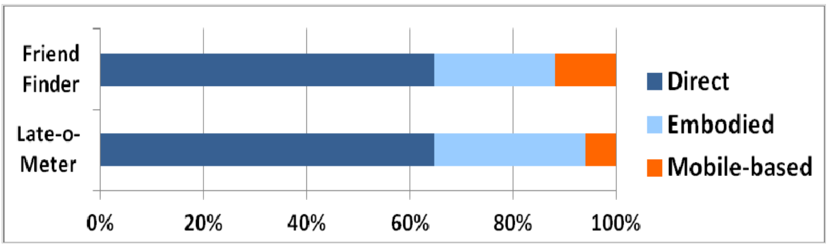


Figure 5.11. Distribution of preferences: Navigation phase.

Direct interaction was perceived more controllable than bodily (LoM: t =-3.77, df = 16, p = 0.0017) and mobile-based (FF: t = 4.24, df = 16, p = 0.0006; LoM: t = -3.77, df = 16, p = 0.0017) techniques. Moreover, the navigation was seen equally easy, precise, and reliable to browse through linearly arranged objects (LoM) as well as irregularly arranged objects (FF). This quality was seen important since many interfaces contain a mix of linear and arbitrary arrangements. Therefore, the user can keep the same interaction style while browsing through differently arranged items.

Bodily technique was also found equally easy and comfortable for browsing through linearly and arbitrary arranged items. The main disadvantage of the technique, however, was its perceived unreliability. Although MS Kinect recognition worked precisely, the participants mentioned they would not trust the system. Bodily navigation was perceived less trustworthy than mobile-based (LoM: t = -2.1, df = 16, p = 0.056) and direct navigation (FF: t = 1.8, df = 16, p = 0.089).

Mobile-based interaction was found more controllable than bodily navigation (FF: $t = 3.27$, $df = 16$, $p = 0.0048$; LoM: $t = -3.4$, $df = 16$, $p = 0.003$). Mobile-based technique was also found more privacy protective than direct technique (LoM: $t = -2.3$, $df = 16$, $p = 0.034$).

Collecting Results Phase

The preferences of 62% of the participants in saving results phase were given to mobile-based interaction, 24% of participants gave their preferences to bodily technique, and 14% preferred direct interaction (see Fig. 5.12).

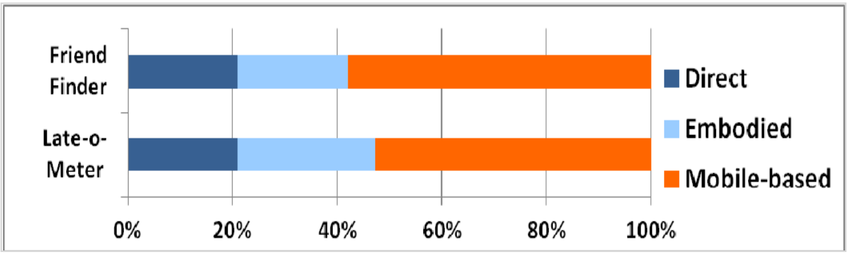


Figure 5.52. Distribution of preferences: Collecting Results.

Mobile-based interaction was found more controllable (FF: $t = -2.1$, $df = 16$, $p = 0.046$; LoM: $t = -4.2$, $df = 16$, $p = 0.0006$) and more reliable (FF: $t = -2.07$, $df = 16$, $p = 0.05$; LoM: $t = -2.5$, $df = 16$, $p = 0.024$) than bodily interaction. Moreover, it was also found more reliable than direct technique (LoM: $t = 2.3$, $df = 16$, $p = 0.03$). The participants emphasized the perceived control, trust, and security of the mobile-based technique: "*I have to be sure I am saving the right thing… I don't want to save something else*".

Bodily interaction was found less comfortable than mobile-based interaction (FF: $t = 2.1$, $df = 16$, $p = 0.05$; LoM: $t = 2.1$, $df = 16$, $p = 0.05$) and direct interaction (LoM: $t = -2.5$, $df = 16$, $p = 0.05$).

5.1.7 Design for Interaction with Personalized Display

Below the author summarize our observations and the participants' comments into a set of recommendations for interaction design on personalized displays.

Keep Interaction Discrete

The users of personalized displays prefer not to demonstrate their interaction on public. This recommendation to personalized displays is in line with existing design recommendations to public displays in passing-by situations as in [152].

The *mobile-based technique* gives users the opportunity to remain discrete. In our experiment, participants mentioned that they do appreciate to remain unnoticed. Mobile interaction enabled users to control the display from any distance, thus letting the users choose any "safe" place.

The *direct interaction*, on the contrary, reveals the fact of interaction. The necessity to interact standing right in front of the public screen made interaction completely noticeable, whereas the users preferred to keep it discrete.

These findings deviate from the user preferences identified for the scenarios that don't involve personalized data. For instance, where people interact with real-life physical objects, they prefer direct and bodily techniques [184]. Therefore, mobile interaction is perceived much less comfortable due to the indirect way of addressing physical objects.

In the experiment, the desire to stay unnoticed was emphasized mostly for the identification phase. However, in the other phases (navigation and collecting results) the discrete interaction did not seem to be that critical. Once the personalized data is put onto the large screen, the users tend to be less concerned about the protection of their identity.

In the navigation phase, the *bodily interaction* was also criticized for making users too noticeable. However, the participants were rather concerned to look ridiculous in public, when using gestural interaction. The concerns were mostly expressed by the people unconfident of the Kinect gestural interaction. They were afraid of confusion, especially in public, if the gestures were not recognized correctly.

Minimize Physical Effort

The users of personalized displays generally prefer to minimize physical motion. This recommendation repeats the guidelines for other interaction scenarios, such as mobile interaction with physical objects [184].

Projected onto the domain of personalized displays, this requirement additionally mimics the users' privacy concerns. Active physical interaction not only attracts attention of the public, but also may slow down the control over the displayed private information. The control, however, is required to be prompt and easy to perform, in the case of potential privacy threat.

The excess in physical motion refers to bodily and direct techniques. Mobile interaction, in all phases, required almost no physical effort.

The *direct interaction* presumed user position next to the large screen. Such position unavoidably caused some physical effort: in order to reach the markers, the users had to move in front of the screen and stretch their arms.

The *bodily interaction* technique, in all phases, presumed some physical motion: crossing the proximity zone for identification, hand gestures for navigation and collecting results. Some participants indeed appreciated such intuitive physical interaction style. Apart from the fun factor, they found it advantageous to be able to interact on the spot without a need in any assisting devices. However, many participants have seen the physical motion rather negatively.

Important to mention, many participants commented that bodily interaction is more entertaining than the other techniques. For example, a female participant noticed: "*I have never tried it, but it's really fun!*" The author supposes that the ratio of supporters of the bodily technique may vary depending on the application character. The applications used in the experiment are rather aimed at *utility* and quick and efficient usage. Therefore, the entertaining nature of the bodily technique might be unnecessary. However, in the applications aimed to be a *toy* for the public the bodily technique might fit better than the other techniques [156, 148].

Provide Position Freedom

The users prefer to be flexible in the choice of their position. Generally, they want to keep some distance to the large screen. This recommendation is specific to the scenario of a personalized display: it addresses the user necessity to stay unnoticed and assures a sufficient overview of the screen.

Participants mentioned that the main advantage of the *mobile technique* is the possibility to control the public display from any distance. Such freedom enabled the users to choose a comfortable position in a given public place; it gave them a chance to have an unnoticeable interaction and also not to move a lot.

Navigation with the *bodily interaction* technique, using hand pointing, was also appreciated for giving participants a relatively flexible choice to position themselves. Unlike the direct interaction the users could choose a comfortable distance from the large screen to ease the browsing process.

Enable Sufficient Display Overview

The users require a constant visual control over the entire surface of a personalized screen. The recommendation maps to the visibility guidelines that are generally applicable to interactive systems [153]. However, in a scenario with personalized displays; it additionally maps to the need of easy and comfortable screen control.

The *direct interaction* technique was often criticized for the necessity to stay too close to the large screen. Apart from already mentioned disadvantages (making interaction noticeable, causing physical effort), such position also hindered the overview of the entire screen. This inconvenience was especially critical in the navigation phase. The participants often had to step back in order to see where the necessary item is located.

The *bodily interaction* technique, on the contrary, made the users interact from some distance, staying centered in front of the display. Such position though made users noticeable, but at the same time maintained a sufficient overview of the entire content. The participants, however, were suspicious about using the bodily interactions in public places. The Kinect recognition is not resistant to occlusions, which often occur in crowded public places.

The *mobile technique* enabled interaction from any position. Thus the user could take care of a comfortable place providing sufficient screen overview.

Exclude Unintended Log-In

The system must initialize the personalization (log-in) only by an explicit command of the user. This recommendation may be mapped to the general usability guidelines, informing designers to avoid erroneous system behavior [151]. However, in the domain of personalized displays the need to avoid an erroneous log-in becomes especially critical. Log-in is a highly sensitive moment in the process of interaction with personalized display. Therefore, users need to be familiar and confident with the technique that triggers the log-in. The *mobile* and *direct interaction* techniques assured that the log-in process entirely controlled by the user's input.

The *bodily interaction* technique, however, exploited rather unusual technique: proximity-based identification inspired by the metaphor of coming closer [76]. Although the idea was generally appreciated by the participants, they saw its main drawback in insufficient visibility. The proximity-based identification assumes that the user remembers how the system works, and where the proximity border lies. Such an invisible or barely notable border may easily cause an unintended identification: the user just comes closer to the display and suddenly the data is shown! In the narrow places, such as office passages, the user may personalize the screen without intention just by passing by the display. Therefore, the proximity border should be clearly indicated, for instance, explicitly drawn on the floor.

Avoid Forgetting to Log-Off

The system must provide the means to remind the user about logging-off or perform the log-off automatically. This recommendation reflects the specific need for users when interacting with personalized data. The users need to be sure to remove the personalized data from the public display when interaction is finished.

The positive side of the *proximity-based identification* was the convenient and intuitive way to log-off. The participants emphasized that the bodily interaction technique was especially beneficial for logging-off. The automatic removal of the personal data is logical and desired when the user abandons the display. It could also secure the cases when the user forgets to log-off.

Allow Quick Exit

This recommendation, specific to personalized displays scenario, supports the user feeling of secure interaction. In order to feel secure, the user needs a leverage to remove the personal data from the public display as quickly as possible.

The *mobile-based interaction* technique enabled the users to react quickly to the surrounding context, and hide the personal data with a single tap. The participants mentioned it as an important aspect. For example, one of them commented: "*I can immediately say: on or off*".

The identification phase of the *bodily interaction* also allowed a quick reaction to the situation: in case of an observation danger the user can quickly step back from the display.

However, the participants emphasized that they need to sufficiently trust the system in order to be sure the system will react in an appropriate way. Since the technique was rather unusual for the majority of the participants, they were suspicious about its reliability.

In the collecting results phase, the participants also appreciated the quick completion of the final action. The *bodily technique* in this phase had a certain disadvantage: the performance of a gesture made the saving slower compared to the other techniques.

Exploit Well-Known Metaphors

The usage of well-known or real-life metaphors can be considered as a general recommendation to interactive systems [153]. The metaphors are widely applied in different scenarios, for instance, in the design of gestural interactions between multiple displays [151].

In our experiment, *mobile-based interaction* with the personalized screen was understood and mastered quickly due to the replication of learnt metaphors from the everyday life. A mobile device was often associated with a joystick or a remote control. The employment of the metaphor enabled people to easily learn and understand the technique: "*It reminds me of a TV remote control*".

Due to the employed metaphors, *bodily interaction*, for identification and navigation, was also found intuitive. To initialize the interaction people had to come closer, just as the proximity-based log-in. The conversation is closed once the person leaves the counterpart (proximity-based log-out). To highlight a distant object, people point at it (gestural pointing).

In collecting results phase, participants also emphasized the importance of metaphors. Thus, the mobile-based saving supported the real-life metaphor of "taking it with me", by saving the result to the mobile "depot" which always accompanies the user.

In the direct interaction technique, the metaphor of "taking it with me" was even stronger, since it was supplemented with the physical action: the user took the object directly from the screen.

Ease "Search and Act" Process

During navigation, the "search" part of the selection process and the selection "act" itself should be supported on the same screen (in case of multiple screens). Moreover, the "search" and "act" parts should be performed within the same interaction flow, e.g. with finger pointing.

This recommendation can also be met in other interaction scenarios that involve several heterogeneous displays. The main interaction issue refers to the uncomfortable focus switch between two screens [184, 31].

In our experiment, *bodily interaction* in navigation phase was appreciated by participants for the possibility to interact from a distance. Apart from a good overview of the screen, such position eased the selection process. Indeed, with the pointing technique, the "search" part of the selection process and the selection "act" itself can be done simultaneously, focusing on the same screen.

Although *mobile-based interaction* was mentioned to be the quickest way to navigate through the content, its main disadvantage was the inconvenient control switch between the small and the large screens. Indeed, the "search" component of the selection process had to be accomplished on the large screen, whereas the "act" component – selection command – using the mobile screen.

5.1.8 Summary

In this work the author explored direct, bodily, and mobile-based techniques for interaction with personalized public displays. Three phases of interaction were examined: identification, navigation, and collecting results. By means of an experiment the author compared the techniques direct, bodily and mobile-based interaction in each phase. The experimental results give an insight into the particularities required for each of the interaction techniques when used for with personalized public displays. The results show that the requirements to the interaction with personalized displays differ from other interaction domains. The work is rounded up with design recommendations derived from the experimental results. The recommendations summarize critical aspects that should be considered when designing interaction for personalized displays.

5.2 Merging Physical and Mobile Interaction on NFC-enhanced Public Displays

The following work studies techniques of mobile interaction for multi-user scenarios at public display. Two types of mobile interaction are explored on an NFC-enhanced display: View-on-Mobile and View-on-PD. Each technique involves both public and mobile display into the presentation of personalized data; however, the degree of involvement of each display varies. By means of an experiment, the author investigates how the presented techniques are perceived in different scenarios, in terms of controllability, reliability, privacy protection, and comfort. The analysis of the experiment provides insights into the factors which drive user preferences in interaction techniques.

5.2.1 Interaction with Personalized Public Displays

Despite the sensitivity of a public scenario, the use of personalized displays has become increasingly common [128]. The vivid interest to personalized displays can be observed not only in scientific society [38, 211], but also in the real-life projects, such as Interactive Video Wall in Copenhagen [94] or CityWall display in Helsinki [165]. These examples show that in spite of the awareness of privacy issues that can result from sharing the personalized content [38, 85, 118], people do place their private data on public displays.

One of the main challenges for the designers of personalized displays is to provide *trustworthy* interaction with the personalized data [69, 181]. Research works on trust issues derive design factors which are critical to maintain user trust: *control*, *reliability*, and *transparency* of interaction and system actions [70, 18, 124]. In the domain of personalized displays, these factors become even more crucial. Users need an ultimate control over their personal data exposed on a public screen. The means for the control should be reliable and available any time. Finally, the users need a clear understanding what actions the system performs with their data, and how the system will react on their input.

Comfort of interaction is another aspect, critical for user acceptance of a personalized display [34]. Interaction with personalized data should be fluent and easy to perform. Thus, in case of potential discomfort, users can quickly react to the situation, for instance, by quickly removing or occluding the private content.

The user discomfort in the scenario of a personalized display can be caused by the privacy level of data [181] and the social context [111]. Therefore, when designing interaction with personalized displays, it is critical to study how interaction techniques are perceived in various contextual situations.

Physical interaction is enabled by physical proximity between an interactive tool and the display. The interactive tool can be the user's hand, finger or an assisting device, such as a pen [175] or an NFC-enabled mobile phone [31]. The technique exploits a real-life metaphor of "touching": a person activates an object by touching it. The touch interaction can be enabled by means of touch-surfaces or by other technologies, such as a matrix of NFC tags [196, 31] or camera-based hand recognition [194, 76]. Studies on direct interaction show that users perceive the technique as natural, fast, reliable, enjoyable, and easy [184, 32]. The direct manipulation of the objects on a large screen supports the feeling of control. Therefore, direct technique is suitable for and frequently used on personalized displays [76, 94, 165]. Its only disadvantage is the tendency to negatively affect user's privacy. By manipulating the data right at the screen, users uncover the fact of interaction and the data explicitly belonging to the user [85].

Remote mobile interaction is supported by a mobile interface which is specially tailored to control the public display. The mobile interface can provide tools for the control (e.g. buttons) or it can replicate the entire public display in miniature size [24]. The advantage of the mobile interaction is the possibility to interact from any distance. Thus, users can interact unnoticeably and maintain their feeling of privacy. Moreover, mobile interface enables users to react quickly and reliably to the changing surrounding situation. The disadvantage of mobile technique is the necessity to control two heterogeneous screens at the same time. The unavoidable visual focus switch is generally perceived as inconvenient [184, 43, 169].

Since both types of interaction show distinct advantages, in the current work the author decided to merge them.

5.2.2 Merging Physical and Mobile Interaction

In order to combine direct and mobile interaction, the author developed an NFC-enhanced public display. Users could interact with the display by touching the NFC tags attached to the screen, using NFC-enabled mobile devices.

Below the author gives an overview of the research works which inspired the idea of the NFC-enhanced screen. Then, the author provides details on the screen implementation and describes the merged interaction techniques.

NFC-augmented Displays and Objects

The invention of NFC technology brought to the research world lots of ideas for augmentation of physical objects and interactive displays [166, 212]. This style of interaction inherits the properties of mobile and direct interaction. The touch of an NFC-tag enables the users to augment a physical object or a display, using the mobile screen.

The idea to augment an interactive display using mobile phone, indeed, originated decades ago. Thus, back to 1993 Fitzmaurice presented the idea of augmenting a large display using a mobile device [59]. Their prototype presented a large digital map which showed contextual information on a mobile display.

Augmentation of a digital display by means of NFC tags was presented by Rukzio and colleagues [196]. The authors introduced an NFC-enhanced laptop monitor. The matrix of NFC tags was attached to the back side of the laptop screen. Users could interact with the content on the front side of the screen by means of their NFC-phones. The touch of a monitor region was registered by the matrix element on the back side and triggered the reaction on the screen. The NFC-augmented monitor showed to facilitate the transfer task: Using the technique people could perform the transfer significantly faster compared to the conventional mobile menu approach.

Static displays, such as paper posters or maps can also be augmented by NFC-tags. Reilly and colleagues attached a matrix of NFC tags to the back side of a paper map [174]. Touching map regions with an NFC-enabled mobile phone, users could retrieve augmented information about the map objects. Broll and Hansen presented a city poster augmented with multiple NFC tags. The tags facilitated the tasks frequently performed by the citizens, such as exploring new locations in the area [30]. The NFC-enhanced posters showed to be more enjoyable and more efficient in use, compared to the conventional task execution with a mobile phone.

Dynamic NFC-enhanced displays were so far developed only in the form of projected displays. Broll and colleagues presented a display projected on a matrix of NFC tags attached to a wall [31]. Users were interacting with the items of the projection, activating thus the respective tags. Since the touch event was registered by the server, the reaction of the system was dynamically updating both, mobile and large projected screen. A similar technology was presented by Hardy and Rukzio [81]. The authors studied various interaction techniques on a display projected on a matrix of NFC tags. Although projected NFC-enhanced displays use the full potential of public and mobile screens, they suffer from the occlusion problems. Mobile devices occlude tags and display items, thus making selection process more difficult.

All in all, the NFC-enhancement shows a great potential for augmentation of paper posters, maps, digital and projected displays.

This work presents an augmentation of a large public display where the tags are placed on the front screen. The advantage of this approach is its flexibility: the removable grid of NFC-tags can be attached to any display. The disposition of the tags can be tailored to an individual content.

Enhancing a Public Display with NFC-tags

The NFC-enhanced screen was implemented on a large non-touchable display. The grid of NFC-tags was attached to the front side of the screen. In order to better address the screen content, the grid of tags was arranged not in a rectangular order of rows and columns, but in an order matching the content of the application, the map of Europe (see Fig. 5.13).

Figure 5.13. Grid of NFC-tags attached to the screen.

The screen of the display was first covered by a transparent adhesion foil. Such foils are used as a protection layer for different kinds of displays. It can be attached without glue directly to the surface of the display and removed any time without leaving any traces on the display surface.

Figure 5.14. Grid of NFC-tags integrated into map content.

The NFC-tags were attached to the transparent foil. Since the foil was absolutely transparent, the content of the screen remained perfectly visible. The NFC-tags were harmonically integrated into the display content: having a decent white colour they were not standing out of the Europe map. Being small and round, the tags were used as markers attached to European cities (see Fig. 5.15).

The users could interact with the NFC-enhanced screen by means of a mobile phone. In order to select a city on the European map they had to touch the corresponding tag with an NFC-enabled phone. Each tag of the matrix contained a unique identifier, written on the tag in advance. Once the touch event has been registered by the phone, the phone immediately informed the server. The server received information about the tag identifier and the identifier of the phone which performed the touch. Having this data, the server extracted the position of the tag within the matrix. Since the matrix elements had fixed positions on the screen, the server could initiate the appropriate update of the large screen and the mobile screen.

As a result, the system could promptly react to the touch events, dynamically changing the content of the large and mobile displays. Figure 5.15 schematically illustrates the communication between the mobile device, the server, and the large display.

Figure 5.15. Communication between NFC matrix, mobile device, and server.

The NFC-enhanced display was implemented on a large display of 42 inches in diagonal with metal-isolated NFC tags of 3cm in diameter. The tags can hold from 48 bytes to 2 Kbytes of data. The mobile communication was supported by Android-based Samsung Nexus devices.

5.2.3 Interaction Techniques

Using the NFC-enhanced public display, the author implemented two interaction techniques. The main goal of the derived techniques was to exploit the advantages of mobile and direct techniques and to overcome their main issues: the uncomfortable focus switch and the exposure of personalized information in public.

The first technique, View-on-PD (PD stands for Public Display) focused more on interaction comfort. It utilized mostly the surface of the large display, placing on mobile screen only the secondary information. The technique aimed to solve the problem of the focus switch: presenting necessary information mainly on the public screen, it minimized the need to work with the mobile screen.

Once the user has activated a tag on the large screen, information related to the tag appeared on the public screen, right next to the tag (see Fig. 5.16 left). Since the information appeared in a small card next to the user, the technique also provided privacy protection: the card could be easily occluded by the user body. Of course, an observer standing next to the display, such as another user, could still see the card content.

The second technique, View-on-Mobile focused more on user privacy. Once the user touched a tag on the large screen, the information card appeared on the mobile screen (see Fig. 5.16 right). The mobile screen extended the content of the large display, utilizing the metaphor of a keyhole. A user could augment the surface of the large screen with a small mobile window. If the users were not comfortable watching the information standing next to the large display, the technique enabled them to "take away" the information card, carrying it on the mobile phone to any comfortable distance.

The touch of a tag in both techniques was accompanied by a decent sound signal, to provide a feedback of a successful connection.

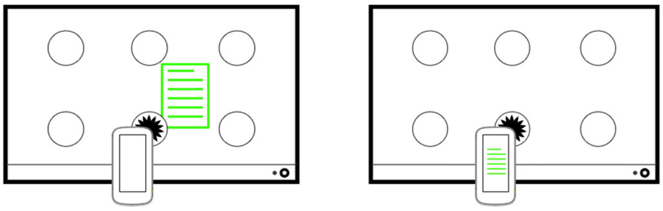


Figure 5.16. View-on-PD (left) and View-on-Mobile (right) techniques.

5.2.4 Experiment

The critical contexts in the scenario of a personalized display are the *privacy level of data* and the surrounding *social context*. By means of an experiment, the author aimed to understand how users perceive the presented interaction techniques in different contextual situations.

In particular, four situations were analyzed:

- Single user interacts with neutral data
- Multiple users interact with neutral data
- Single user interacts with private data
- Multiple users interact with private data

Single user referred to the situation when the user was interacting with the display alone, inside a university public area. Multiple users referred to the situation when several persons performed independent tasks on the same display. Important to emphasize, the author did not analyze collaborative tasks, but only independent tasks.

In either interaction technique in each contextual situation the author wanted to capture the user perception of interaction transparency, controllability, comfort, privacy protection, reliability, and trust. Moreover, the author aimed to understand which technique the users subjectively prefer in each situation and what drives their preferences.

Travel Planner

The experiment was conducted with a sample application called Travel Planner. The application was designed by the students and aimed to support their peers who want to travel around Europe, but have a limited budget. The applications helped students find destinations that would match their interests and would fit to their personal budget. The students thus could optimize their travel time and expenses.

The large display showed the map of central Europe. The main cities on the map were marked by NFC-tags attached to the cities. By touching a tag with the mobile phone, users retrieved information about the respective city. The information card contained the data on the city name, population, size, transportation means, recommended time of stay, and estimated expenses. The user then could add the city to the personal trip or close the card. The personal trip with the list of the selected cities was saved on the mobile phone and could be retrieved any time.

The display supported the interaction with the map in two ways: with View-on-PD or View-on-Mobile techniques.

View-on-PD

Once the user has touched a tag, the information card appeared on the public screen, right next to the activated tag (see Fig. 5.17). The user could add the city by touching the tag again. The advantage of the technique was in the possibility to open several cards at the same time, for instance, to compare the prices. The opened cards stayed on the screen about 20 seconds. If the cards were occluding each other, the latest opened card was placed on the top.

View-on-Mobile

Once the user has touched a tag, the information card appeared on the mobile display (see Fig. 5.18). The user could add the city to the trip by means of a button provided on the mobile screen. The advantage of the technique is the possibility to "take away" the card from the large screen and watch it on the mobile phone at any position, not necessarily standing next to the public screen.

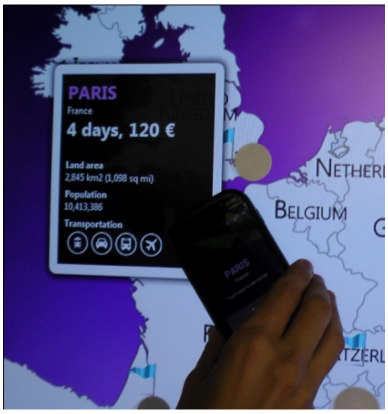


Figure 5.17. View-on-PD technique.

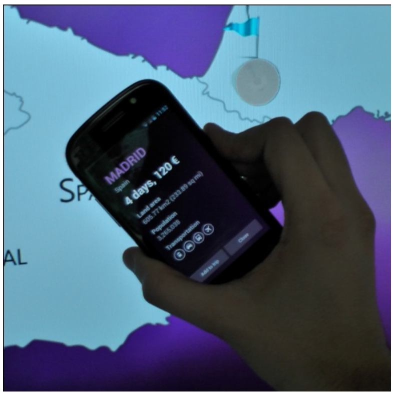


Figure 5.18. View-on-Mobile technique.

Single and Multiple Users

Both techniques supported interaction of single and multiple users. With View-on-PD technique, several users would share the large screen to retrieve their cards (see Fig. 5.19 top); with View-on-Mobile technique the cards would appear on individual mobile phones (see Fig. 5.19 bottom).

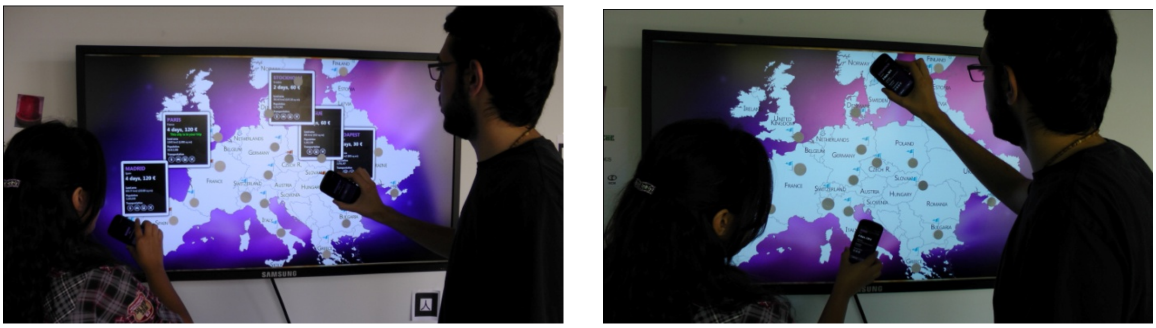


Figure 5.19. Multiple users interacting with View-on-PD (top) and View-on-Mobile (bottom) techniques.

Privacy level of information

Travel Planner supported two modes: neutral and budget-aware. In the neutral mode, the openeded information cards contained only *neutral data*: the city information and estimated budget for the visit. The mode provided no link to the personal budget of the users. The users could browse the cities familiarizing themselves with the city information, and estimated expenses.

In the budget-aware mode, the opened cards were linked to the user's limited budget, exposing their *personalized data*. An opening card was informing the user not only about the general city data, but also whether the city visit can fit to the user's budget. The budget limit was saved on the user mobile device. In this mode, the frames of the information cards were displayed in red or green. The red colour informed the user that the city visit is out of the budget; the green colour showed that the visit fits to the budget. Additionally, the cards contained a message, printed in red or green, informing the user about the exact amount of money remaining in the trip budget. The design of information cards in View-on-PD technique is illustrated in the Figure 5.20. The design of View-on-Mobile cards was done similarly: the coloured frames were added to the mobile screens, and the red or green messages integrated into the mobile card.

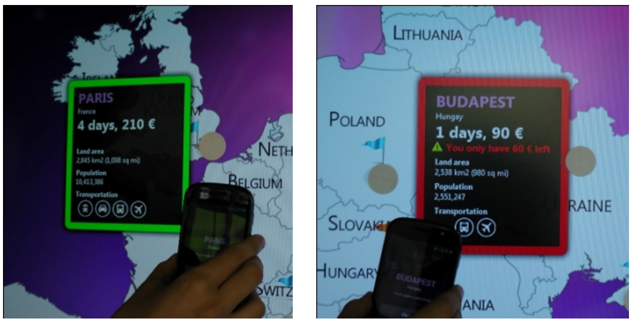


Figure 5.20. Information cards in budget-aware mode.

To summarize, Travel Planner supported both NFC-enabled interaction techniques, View-on-PD and View-on-Mobile. It supported two levels of privacy: neutral and budget-aware mode, and could be used by single or multiple users. Therefore, the application was perfectly suited for an experiment with different combinations of context.

Experiment Design

The experiment was conducted in a public place of a university (see Fig. 5.21). The NFC-enhanced screen was mounted at the wall in a passage. The circulation of people in the passage is quite low; therefore it was possible to conduct the experiment under controlled and consistent conditions. In spite of being so quiet, the passage gives a feeling of a public place.

The experiment was conducted individually, following the within-groups design. A moderator and an assistant were present during the experiment. After an introduction of the experiment procedure and the Travel Planer application, the moderator explained the idea of the NFC-enhanced screen. He did not demonstrate, however, how exactly the display can react to the touch event. Then, the participant was asked to perform the tasks. The experiment consisted of eight tasks.

Figure 5.21. Experiment setting: a public area of a university.

The first four tasks were performed in the neutral mode. The participants had to browse through the cities, watch the information cards with neutral data, and find a solution to the given task. The tasks in the neutral mode were asking, for example, to “Find 3 cities, each with a population less than 3 million and add them to your trip” or “Find 3 cities, each with a land area smaller than 500 km^2 and add them to your trip”.

Tasks 1 and 2 were performed by the single user: the participant was interacting alone. One task was performed with View-on-PD technique, the other task – with View-on-Mobile.

Tasks 3 and 4 were performed by multiple users: another user was interacting with the screen simultaneously with the participant. The role of the other user was played by the assistant, our colleague, not well familiar to the test participants. The other user was performing a task similar to the participant: browsing through arbitrary cities. Here again, one task was supported by View-on-PD, and the other – by View-on-Mobile. The order of the techniques was consistent with the first two tasks. Half of the test participants always started with View-on-PD, and half started with View-on-Mobile.

The last four tasks were performed in the budget-aware mode, showing data on users’ private budget limit. The tasks were formulated, for example, like this: “You have a budget of 300 Euros. Find at least 4 cities that can fit into your budget and add them to your trip.” or “You have a budget of 350 Euros. Plan a trip that spans at least 10 days”. The participants had to browse through the cities in order to find the solution. The author made sure that every task had a solution. However, programmatically the author assured that each participant browses at least through four green cards and bumps into at least three red cards, before finding the solution (see Fig. 5.22).

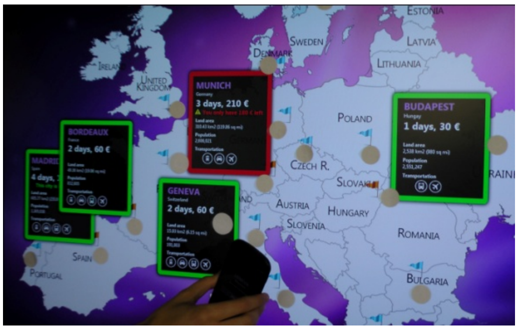


Figure 5.22. Displaying budget-aware cards with View-on-PD technique.

Tasks 5 and 6 were performed by single users, participants being alone. One task was supported by View-on-PD technique, and the other – by View-on-Mobile technique.

Tasks 7 and 8 were performed by multiple users, the participants in the presence of another user. Another user, played by the assistant, was performing similar tasks: opening arbitrary cards with private data. The cards of the assistant were linked to his predefined budget and were also appearing in red or green.

One task was supported by View-on-PD, and the other – by View-on-Mobile technique. The order of technique was kept consistent throughout the experiment.

After each task participants the author asked to fill in a questionnaire. The questionnaire aimed to capture how either technique is perceived in the given context, in terms of transparency, controllability, comfort of use, reliability, privacy protection, and trust. The questionnaire was based on the studies focused on comparison of different interaction techniques [184, 32, 15, 152].

The questions were formulated as statements. The participants had to rate the statements on a 5-Likert scale, from "strongly agree" to "strongly disagree".

Q1: *"The display reacted as I expected"*

Q2: *"I had control over the system"*

Q3: *"I found the system comfortable to use"*

Q4: *"The display protected my privacy in appropriate way"*

Q5: *"I found the system reliable"*

Q6: *"I found the system to be trustworthy"*

Additionally to the questionnaires, the participants were asked to express their preferences for each contextual situation. For this purpose, after each combination of context, the author asked the participants about their preferences. To express the preferences, the participants had to rate the statements on a 5-Likert scale from "strongly agree" to "strongly disagree". The statements were capturing user satisfaction with either technique, in the given context. For example, after the tasks 5 and 6 (private data, single user) the preference questions were formulated like this:

"When viewing budget-aware cards alone, I prefer to view the cards on public display"

"When viewing budget-aware cards alone, I prefer to view the cards on mobile display".

Table 5.2 gives an overview of the study procedure.

Apart from the order of the interaction techniques (which was counterbalanced between the participants), the order of the tasks was fixed. The neutral mode was preceding the private, budget-aware mode; the tasks for single user always preceded the tasks for multiple users. The fixed order was chosen with purpose. Since the author aimed to capture the user perception of privacy and trust, it was important to see how people perceive the techniques first in a less critical situation, being alone. The experiences of interaction with others could influence this perception and result in higher estimations, just due to a more "relaxed" context. Giving the "more relaxed" context first reduced the chance of getting the biased estimations, based not on the real feelings but on comparison with more critical context. The same rationales were followed in fixing the order of neutral and private modes.

Task	Context	Technique
1	Neutral data & Single user	View-on-PD
2	Neutral data & Single user	View-on-Mobile
	Questions on preferences: Neutral data & Single user	
3	Neutral data & Multiple users	View-on-PD
4	Neutral data & Multiple users	View-on-Mobile
	Questions on preferences: Neutral data & Multiple users	
5	Private data & Single user	View-on-PD
6	Private data & Single user	View-on Mobile
	Questions on preferences: Personalized data & Single user	
7	Private data & Multiple users	View-on-PD
8	Private data & Multiple users	View-on Mobile
	Questions on preferences: Personalized data & Multiple users	

Table 5.2. Experiment procedure.

Participants

Twenty persons participated in the study, among them 8 female and 12 male. The participants were aged from 20 to 31 (average of 24) years, coming from Germany, Egypt, Canada, Malaysia, New Zealand, and South Korea. The participants had background in IT, Engineering, Law, and Economy. Most of them were exchange students who came to the university for a year or several months. The remaining participants were locals, but having experience as an exchange student. The problem of travelling with a limited budget around Europe was familiar to all of the participants.

5.2.5 Experiment Results

The results of the questionnaires were analyzed statistically, using repeated measures ANOVA test. Below the author provides the results on each design factor and on the preferences given by the participants in each contextual situation.

Transparency. No significant differences between two interaction techniques were found for the category transparency, $F(3.1, 59) = 3.9$, $p > .05$. The participants quickly understood how the techniques worked and estimated the rankings for the question Q1 ("The display reacted as I expected") similarly highly in all contextual scenarios (M: from 4,5 to 4,75).

Reliability. Similarly, no significant differences were fond for the perception of reliability, $F(4.6, 87) = 2.4$, $p > .05$. Since the techniques were functioning in a consistent and stable way, the question Q5 ("I found the system reliable") also received high estimations, independently on the context (M: from 4,15 to 4,35). The participants perceived the reliability rather tightly related to the technical functionality, and not control over privacy.

Controllability. In the scenario of multiple users interacting with neutral View-on-Mobile technique was perceived significantly more controllable than View-on-PD technique $F(4.3, 81.6) = 3.6$, $p = 0{,}037$.

Comfort. Similarly, when multiple users interacted with neutral data View-on-Mobile technique was found significantly more comfortable that View-on-PD technique, $F(2.7, 51) = 8.4$, $p = 0{,}001$.

Privacy. Significant differences in privacy perception were found almost in all context combinations. View-on-Mobile technique was estimated more privacy protective than View-on-PD technique when multiple users interacted with neutral data ($F(2.9, 54.4) = 15.1$, $p = 0{,}011$), when single user interacted with private data ($F(2.9, 54.4) = 15.1$, $p = 0{,}004$), and when multiple users interacted with private data ($F(2.9, 54.4) = 15.1$, $p = 0{,}005$).

Trust. View-on-Mobile technique was perceived more trustworthy when multiple users interacted with neutral data $F(3.5, 66.7) = 5.4$, $p = 0{,}012$.

Preferences. The analysis of the preferences revealed that the View-on-Mobile technique was estimated significantly higher almost in all context combinations: when multiple users interacted with neutral data ($F(1,19) = 8.1$, $p = 0.014$), when single user interacted with private data ($F(1,19) = 9.3$, $p = 0.006$), and when multiple users interacted with private data ($F(1,19) = 50.3$, $p < 0{,}0001$).

5.2.6 Discussion

The study showed that the preferences of the users were often driven by various factors: privacy concerns, visual comfort, space conflicts, and collaboration potential.

Privacy Concerns

Privacy issues arise not only from the exposure of highly-private data, but also from the presence of another user. Even if the exposed data has a neutral character, the users feel uncomfortable due to the potential observation. Therefore, users prefer to keep their data – private of neutral – protected from the eyes of an observer.

The technique View-on-Mobile was appreciated for the possibility to keep data on a private depot, their mobile phone, when the feeling of privacy threat was increasing. As a result, View-on-Mobile was perceived as more privacy protective and got user preferences when private data was exposed and when a second user was present.

Visual Comfort

Comfort of interaction was another driver motivating user preferences. Partially, it was related to the overview of the information. Although View-on-PD technique gave users the opportunity to work with only one screen, View-on-Mobile technique was mentioned to be as comfortable.

Indeed, the View-on-Mobile technique successfully solved the problem of *visual focus switch*. In order to touch an NFC-tag, the users held the mobile phone in vertical position. They oriented the body of the mobile device in the same plain with the large screen. Since the touch had to be performed from a short distance, the surface of the mobile display was almost coincided with the surface of the large screen. Interacting with the tag on the large screen, user got information card on the mobile screen right next to the tag, exactly at the location of the user's current visual focus. Thus, the focus switch problem was resolved.

Since both techniques were giving an equally comfortable overview of the data, the users divided their preferences, when interacting alone with neutral data. However, when multiple users interacted with neutral data another critical factor came into play: space conflicts.

Space Conflicts

When users interacted alone, View-on-PD technique enabled them to decide which cards to open or to close. Therefore, they could control the amount of the cards on the screen. When the density of the cards started to be uncomfortable, the users were in control to close unnecessary cards.

However, when another user was interacting simultaneously, the control over the cards was much reduced. The uncomfortable density of the cards might have originated from another user. However, the control over the cards was not there anymore: the unnecessary cards might have belonged to the other user and thus could not be closed. The large amount of cards hindered the overview, irritated the users, and made the execution of the given tasks harder.

View-on-Mobile technique, on the contrary, enabled users to interact with as many cards as they needed. It gave them the ultimate control over the cards. Moreover, they could step away from the display at any comfortable position, not to stand too closely to the second user. Indeed, the technique supported the subconscious need of the people to choose a comfortable distance to the co-interacting person, depending on their familiarity and culture [80].

As a result, in multi-user scenario people perceived View-on-Mobile technique as more comfortable, more controllable, and more trustworthy.

Collaboration Potential

Curiously, the cards opened by the other user sometimes helped the participants to solve the given task. However, when the second user (the assistant) was looking inside the user cards, the participants commented to feel uncomfortable. View-on-Mobile technique protected them from such implicit potential of observation.

To summarize, View-on-Mobile technique optimally addressed the issues typical for interaction with personalized displays. It solved the problem of the focus switch between heterogeneous displays, protected user privacy in appropriate way, and eliminated potential space conflicts. The technique, thus, is recommendable to use when interacting with private data or when multiple users are present.

5.2.7 Summary

The presented work aimed to find which interaction techniques are appropriate for different scenarios of a personalized display. By means of an experiment, two techniques were examined: View-on-PD and View-on-Mobile. The techniques were supported by an NFC-enhanced public screen. The experiment revealed how the techniques are perceived when interaction with the display happens in a critical context. The author analyzed four contextual situations manipulating the privacy level of data and the social context.

5.3 Studying User-defined iPad Gestures for Interaction in Multi-display Environment

The last work in this chapter investigates how social context influences user preferences in gestural (bodily) interaction [113]. The work studies the iPad gestures that users naturally perform for data transfer. The author examines the transfer between two iPads, iPad and a tabletop, and iPad and a public display. Three gesture modalities are investigated: multi-touch gestures, performed using iPad display, spatial gestures, performed by manipulating iPad in 3D space, and direct contact gestures, involving the physical contact of iPad and other device. The author reports how social context impacts user choices of the modalities and gesture types, and derives critical points for the design of iPad gestures.

5.3.1 Motivation

As predicted by Mark Weiser back in 1991 [215], tablet-sized devices, such as iPad, gain increasing importance in the modern world. Among the rich interaction possibilities, these devices support the convenient gestural interaction. Gestural interaction is known to be intuitive, quick, comfortable [106], as well as fun and engaging [16]. However, a great challenge is to design the gestures in the way which would match the way people interact naturally.

The goal of this work is to find the gestures that users naturally perform with iPad[2] for data transfer. By means of an experiment the author studies the gestures in three settings: iPad and iPad, iPad and a tabletop, iPad and a public display. The author analyzes the rationales behind the user choice of the gestures, and derives the guidelines for the gesture design.

5.3.2 Studies on User-defined Gestures

User-defined gestures have been widely studied in the domains of full body interaction [198], tabletop interaction [215, 177, 218], interaction on vertical surfaces [16, 20], as well as mobile interaction [183].

The research on natural gestures has also been done in the domain of multi-device [122, 219] and multi-display [121, 175] environments. For example, Kray et al. [106] studied user-defined mobile gestures people perform for connection tasks between mobile devices, public displays, and tabletops.

However, not much work has been done so far to study the gestures people naturally perform with tablet-sized devices. Rekimoto [175] presented pen-based gestures for data transfer between tablet-sized devices. To enable the transfer, the technique exploits the touch screen of the tablet device. However, apart from the touch screen, a modern tablet device possesses more interaction possibilities enabled e.g. by integrated accelerometer or camera. Therefore, there is a need to understand which gestures are appreciated and preferred by the users.

5.3.3 Modalities of iPad Gestures

Thinking of possible gestures that can be performed on touch surfaces or mobile devices, one can distinguish three gesture types for iPad devices:

- *Multi-touch gestures* are performed on the iPad screen.
- *Spatial gestures* are performed by rotating, tilting, or panning the iPad body.
- *Direct contact gestures* imply physical contact between the iPad body and the body of the other display. The direct touch gestures can also be performed from the distance,

[2] Here and further the author references only to iPad. However, the discussion can be applied to any tablet PC device possessing similar physical properties and sensory equipment as iPad.

imitating the physical contact by e.g. copying the remote display to the display of the tablet device [24].

5.3.4 Studying iPad Gestures

In order to understand which gestures, in which modalities are intuitively chosen by users when performing transfer tasks on multiple displays, the author arranged an experiment.

The experiment involved two transfer tasks, sending and receiving, performed in three settings: *iPad-iPad, iPad-Tabletop, iPad-Public Display*. The users had to perform the transfer of a large and small object in every setting. The order of the settings was

The experiment was structured into two parts. *Part 1* investigated the gestures that people perform spontaneously for transfer in either setting. Here, the participants were free to choose any modality. However, since such gestures may be strongly motivated by the experiences with other devices, they might not reflect the real user preferences. *Part 2* aimed to correct this effect by giving the participants a chance to explore all possible modalities and then to express their preferences.

The experiment was conducted as a within-group test, individually with every participant. After a short introduction, the moderator explained the basic interaction techniques on iPad, avoiding any demonstrations. Then the participants were asked to complete Part1, Part2, to give their preference in gestures and explain the choice.

Important to mention, in *Part 1* two distances were examined in *iPad-Public Display* setting: closer and farther from the display. Thus, the author could imitate the real life settings: often it is not possible to reach the display.

All in all, 20 persons participated in the test, 14 males and 6 females, aged between 21 and 55 (average 28,8), with occupations of a carpenter, a sports teacher, law students, and IT researchers. Ten have experiences in gesture-based entertainment consoles, 13 participants possess touch-enabled mobile devices, six participants have experiences with tablet devices; among them only two possess an iPad; ten participants have never worked with a tabletop.

5.3.4 Results: User-defined Gestures

Among the gestures performed in *Part 1*, 58% included multi-touch gestures, 17% spatial gestures, and 25% direct contact gestures.

iPad-iPad Setting

The multi-touch gesture for sending of large and small objects was performed by 76% of the participants by *dragging the object with the finger from the participant's own iPad in the direction of the other iPad* (see Fig.5.23, left). The receiving was performed by 50% of the persons, similarly: by *dragging the object with the finger, either directly from the other iPad* *or only using the surface of the own iPad*. The participants usually placed their iPad in the same plane with the source iPad, organizing a common surface for dragging.

Spatial gestures for sending were performed by 20% of the participants by *pouring out of the object from the iPad* so that the object “slides” to the other iPad (see Fig. 5.24, left). For receiving, however, spatial gestures were generally found inappropriate and hard to find. Just two participants performed *pulling the iPad from the source iPad in the direction of their body.*

Direct contact gestures for sending were performed by 4% of the participants by *placing or “posting” the object on the other iPad by holding the own iPad over the receiving iPad.* For receiving about 50% of the participants *placed the iPad over the other iPad to copy the object or attract it by imagined gravity* (see Fig. 5.25).

Figure 5.23. Sending the object with multi-touch gestures

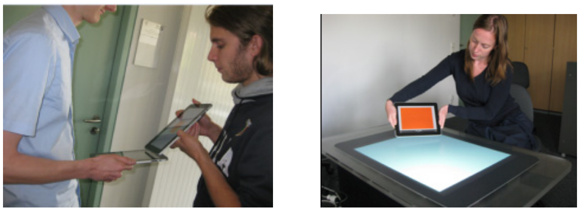


Figure 5.24. Sending the object with spatial gestures

Figure 5.25. Receiving the object with direct contact

Figure 5.26. Receiving the object with multi-touch and direct contact gestures

iPad-Tabletop Setting

For sending the majority of the participants (80%) performed multi-touch gesture: *dragging the object with finger from iPad in the direction of the table* (see Fig. 5.23, right). Just few participants touched the table. For receiving the majority (70%) was *dragging the object from the table to iPad* (see Fig. 5.26, left). Most of participants touched both the table and the iPad.

Spatial gestures for sending were performed by 12% of the participants. They mostly included *pouring the object out from iPad to the table* (see Fig. 5.24, right). Some participants shook the device additionally, to "make sure the object slides down". For receiving almost no one has performed the spatial gestures, commenting that it is hard to find a good match.

Direct contact gestures for sending were performed by 8% of the participants, by *"posting" the object to the table almost placing the device physically on the surface*. For receiving, about 30% of the participants performed *holding the iPad over the table to "capture*" the small object (see Fig. 5.25, right) or *placing the iPad at the table to "imprint"* the large object. Seven participants placed the iPad on the table before interaction.

In both iPad-iPad and iPad-Tabletop settings, for the large objects the participants often used several fingers. The tasks involving the small object were performed mostly with one finger. Moreover, the large objects were frequently scaled down before the transfer.

iPad – Public Display Setting

For sending, the majority of the participants (75%) chose multi-touch modality by *dragging the object with the finger from iPad in the direction of the display*. For receiving, however, multi-touch gestures were found rather inappropriate. Few participants performed *pulling the corner of iPad by multiple dragging motions*, reminding scratching movements, commenting however that the gesture is not really suitable for the transfer.

Spatial gestures for sending were performed by 15% of the participants. The gestures included *intensive jerky "throwing" of the object in horizontal plane* or "*throwing the object" by tilting the iPad from the body to the horizontal position, orthogonal to the display*. For receiving, 20% of the participants performed the *„scooping up" the object from the large display*.

Direct contact gestures for sending were performed by 10% of the persons, by *"posting" or „pinning up" the object on the large display.* For receiving, direct contact was performed by the majority (75%). To receive a small object the participants were *„picking up" or "capturing" the small object from the display* (see Fig. 5.26, right); to receive a large object they *touched an imagined marker at the large display to copy the content.*

The distance influenced only the sending gestures. The further the participant stayed from the display, the less multi-touch gestures and the more spatial gestures were performed. The phenomenon can be explained by the perceived weight of the transferred object. Throwing with the finger gives an object a minor acceleration. Thus, it can "fly" only through a short distance. Throwing using the whole iPad body can produce a stronger acceleration. Therefore, the object can securely "fly" over a longer distance and not to "get lost" on the way.

Part 2 gave the inexperienced participants the chance to try all possible gesture modalities, and perhaps correct their preferences. As a result, the participants expressed more preferences for spatial and direct contact gestures, in iPad-iPad and iPad-Tabletop settings. These modalities were perceived more creative and more enjoyable than multi-touch gestures. Especially, spatial gestures were referred as "more funny to perform". However, in iPad-iPad setting, even well-matching spatial gestures were often perceived as unreliable. In iPad-Tabletop setting, this disadvantage was diminished due to the presence of the table. Placing the iPad on a solid surface, people felt they can better control the transfer.

Direct touch was less popular in iPad-Tabletop setting compared to *Part1*. The physical dimensions of iPad and the table were reported to be too different to find a well-matching metaphor for a direct transfer (in other words to "pack" the object directly to iPad).

To summarize, the participants favoured multi-touch modality almost in all the settings. The exception were only the receiving gestures in *iPad-Public Display* setting which was most frequently performed with direct contact.

5.3.5 Designing iPad Gestures

Below the author emphasizes the aspects important to consider when designing iPad gestures.

Physical Shape

The flat page-like shape of iPad device affects its physical disposition to the other devices, its orientation, and the set of possible spatial gestures.

The physical touch of other devices occurred mostly in *iPad-iPad* setting and especially in *iPad-Tabletop* setting. The flat shape of the iPad body invites to place it on a flat horizontal surface.

The flat shape also restricts the spatial gestures that can be performed with iPad. Thus, the usage of spatial rotations is rather provoked by the distance to the target device. The further the device, the more spatial gestures are performed. Therefore, spatial gestures were popular in *iPad-iPad* and *iPad-Public Display* settings.

Physical shape also influences the iPad orientation: during the transfer iPad is often oriented in the same plane with the display. For example, it is placed boarder to boarder to the other iPad or a tabletop arranging a continuous surface.

Thinking of Flat Metaphors

Both experienced and inexperienced users rely on real-life metaphors when thinking of well-matching gestures. The most common metaphor refers to a tray (or a salver). The object on the iPad screen is associated with a flat physical object lying on the tray. The iPad is usually held with two hands; its display is associated with the top side of the tray. Therefore, iPad is almost never turned around. If iPad is manipulated in 3D space, the final position is always in the horizontal plane, oriented by the "top of the tray". Grabbing the objects from other surfaces is done similarly to the collection of pieces to the tray.

The object on the tray is perceived as a physical one, having its weight and size. Therefore, the larger the object, the more effort the user needs to move it; the more power needs to be applied to throw the object away from the tray. The gestures for the larger objects require stronger acceleration and involve more fingers.

The metaphor of a flat tray is often transformed into a plate metaphor. The participants often performed pouring out gesture reminding pouring out of a liquid from a plate. Here again the size of the object was mapped onto its weight. The plates with the "heavier" objects were additionally shaken, to ensure the object slides down from the plate.

The metaphor of a spade or a shovel was often applied to receiving tasks. The participants tried to scoop up the object from the source display using iPad as a shovel. Interestingly, the metaphor was applied to both horizontal and vertical surfaces.

Privacy Matters

An important aspect in the design of iPad gestures concerns the private space. The gestures involving the touch of another private device (iPad) are seen as an intrusion into the person's private space. This concerns both the finger interaction and the direct contact of the device body. Even the contactless interaction which presumes close proximity of two devices (see e.g. Fig. 5.23, left) was seen slightly privacy critical.

Another concern relates the direct contact gestures as the demonstration of the interaction on public. Direct contact was criticized by potentially drawing public attention and by the impossibility to interact unnoticed from a distance.

5.4 Summary

This chapter presented the studies on how social context influences user choice and preferences in interaction techniques. Since many different techniques nowadays enable interaction with public displays, it is crucial to understand which technique to offer in current contextual situation.

The author investigated how direct, mobile, and bodily techniques are perceived when interacting in a public scenario. Analysing different cases – interaction alone, in presence of spectators or active co-interactors, interacting with neutral or private data, in single- or multi-display environment – the author reports on the findings and analyzes which factors impact user choice of interaction technique.

Chapter 6

Preserving Privacy in Public Display Use

Personalization of content on public displays is likely to cause the disclosure of user's private data. In order to protect the user's privacy, different protection strategies are used, e.g. the private data is hidden, occluded or blurred. Existing systems usually follow a uniform protection strategy, applying it every time a spectator is detected in the display proximity. However, the necessity in privacy protection often depends on the personal relationships between the user and the spectator. This chapter investigates how the relationship context influences user preferences in adaptation strategies. Additionally, the author studies how privacy level of data and the presence of a mobile device influence this preference. The obtained results can guide adaptation designers in creation of more flexible privacy protection mechanisms.

6.1 Motivation

Modern public displays utilize various strategies to protect private data, for instance, occluding, removing, or blurring the private content. As a rule, these techniques are uniformly applied every time a spectator is detected near the display [34, 211]. However, user willingness to expose or hide private content usually depends on the relationship with the spectator. Thus, people may want to demonstrate the content to their friends, but protect it from the eyes of a stranger.

Different protection strategies can be applied to different relationships. Similarly to the trusted groups in social networks (such as Facebook), spectators can be classified into groups. Each group is assigned a specific adaptation strategy, providing stronger or weaker data protection.

This chapter investigates how relationship context impacts user preferences in adaptation strategies. Using two example applications, designed for community public displays, the author analyzes user attitudes towards protection necessity in different scenarios. Besides the relationship context, the author takes a look at additional context sources such as privacy level of content and the presence of a mobile device in the setting.

After an overview of the existing privacy protection techniques, the author motivates the need for relationship-based adaptation. Then, the experiment is described followed by the comments on the obtained results. Finally, the author presents the application of the results in adaptive public displays.

6.2 Privacy Protection on Public Displays

Adaptation aimed at the personalization of content is widely utilized on public displays. For example, it facilitates the selection of relevant news [210], gives the audience more details about the speaker [137], or reminds the user on important information [34, 211, 185].

Besides evident benefits, the personalization brings the risk of privacy disclosure [137]. In order to avoid the disclosure, public displays can adapt to the presence of spectators using various protection strategies. The protection power of these strategies may vary from strong, such as a complete removal of private data from the display, to very low, e.g. doing nothing or just minimizing the size of the private content. All in all, the display reaction aimed at privacy protection can be classified into the following groups, ordered by ascending protection power:

1. *Do Nothing*: the display is insensitive to the presence of spectators. All private data remains on the screen [137].

2. *Minimize*: private content is minimized in size, or changes its transparency [211] or sharpness. Spectators can still view the private data, though the visual access to the data is hampered. The user can continue interaction with the data.

3. ***Mask***: private content on the display is occluded with blinders [181], pixelized [27], or covered with some neutral display elements [181]. Spectators can clearly see that some content is hidden; however, they cannot view the content itself. Since the private data is protected, users need to interrupt the interaction.

4. *Remove Private Part*: if the personalized content is partially neutral (e.g. selected news) and partially private (e.g. user's name), protection mechanism removes only the private part from the display. Spectators cannot notice that some content is hidden or missing [197]. Users have to interrupt the interaction.

5. *Remove All*: All personalized content is removed from the public screen. Spectators cannot notice that the content is hidden; user interrupts the interaction.

Existing adaptive displays usually follow a uniform protection strategy independently on the personality of the spectator. Such a uniform privacy protection brings certain inflexibility into the system, since the concept of privacy depends on the relationship with the spectator [f21, 92]. People's need to differentiate between trusted groups can be clearly seen on the example of social networks [163]. In online network communities, users specify unique policies for the groups of friends, family, colleagues, etc. and thus control the access to their private data [99]. The need in such relationship-based trusted groups also holds for public displays. The system should be able to determine the group of the current spectators and perform the adaptation according to the group's policy. Such relationship-based adaptation will not only increase the comfort of interaction, but also will increase user trust [130].

Although modern researchers often emphasize the need for a flexible relationship-based adaptation [61, 92, 163], there is no research in the domain of public displays that studies this question in greater details.

Besides the relationship context, other factors influence the necessity in data protection: for example, privacy level of content [34] and the current technical setting [186]. Modern research proposes diverse technical settings for interaction with public displays, starting from mobile devices [186], tablet-PCs, and finishing with AR-helms and stereoscopic glasses [27]. Since current work focuses on the settings available and familiar to a wide range of users, two typical settings are considered: only a public display (*PD-only*) and a public display assisted by a mobile device (*PD-mobile*). Other contexts, such as user activity, current task, etc. might also influence the protection necessity. However, they are often task-dependent and thus their impact is hard to generalize.

6.3 Studying the Relationship-based Adaptation

To summarize, this chapter tackles the following questions:

- How does the relationship with the spectator influence user choice of protection strategy on public displays?
- How does privacy level of content influence this choice?
- How does presence of a mobile device in the setting influence this choice?

6.3.1 Prototypes

In order to address these questions, the author arranged an experiment with two public display applications. The applications, Friend Finder and Media Wall, represent typical content for community public displays: support of a social network and a media gallery. Examples of these content types can be frequently encountered in research works [44, 91, 85, 73, 165], as well as in real-life projects, such as Interactive Video Wall in Copenhagen [94] or CityWall in Helsinki [165]. The examples show that despite the awareness of privacy issues [85, 118] caused by the personalized content, people do place their private data on public displays and do need to protect it.

Friend Finder visualizes a user's social network overlaid over a local map (see Fig.6.1). The application was introduced in the Section 5.1. The users can browse through their peers and retrieve the directions to them. The public display can be operated via a mobile phone client or by means of gestures. An earlier conducted study [109] revealed that the peers' names, pictures, and locations are considered as privacy-critical data. Therefore, a privacy protection mechanism was integrated into Friend Finder: the display executed a uniform masking of peers' pictures and names once a spectator was detected in the proximity [107]. However, intermediate evaluations uncovered the need for a more flexible adaptation: users felt a need to protect their data only from certain individuals.

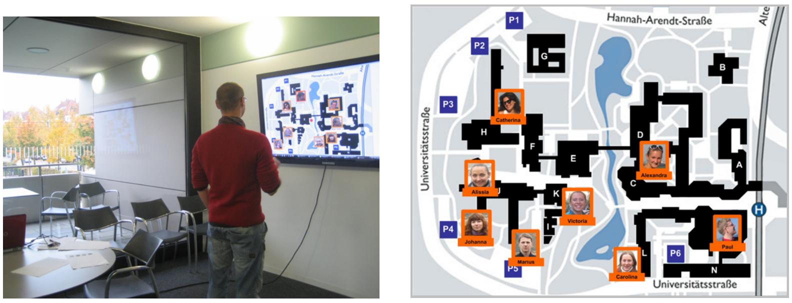


Figure 6.1. Friend Finder visualizes user's social network on a public display.

Media Wall presents a collection of media shared by community members. Users can upload and edit their media in the working space (see Fig. 6.2), view and rank the media of the others. Since privacy concerns are likely to arise when uploading the private media [85, 22], an adaptive protection was required. The need in protection depended on the privacy level of media content and on the spectator personality.

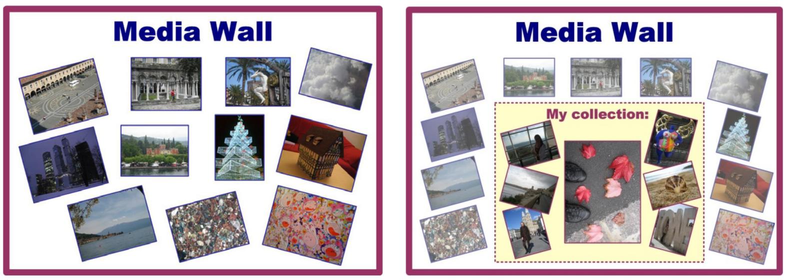


Figure 6.2. Media Wall: start screen (left) and personalized working space (right)

For the experiment the author created several prototypes of each system which differed by the privacy levels and by settings. For Friend Finder, two levels of privacy were provided. The higher privacy version (Friend Finder 2) showed peers' names and portrait pictures (see Fig. 6.1, right). The lower privacy version (Friend Finder 1) showed only names and uniform icons instead of the pictures. For Media Wall, three levels of privacy were created. The low privacy version (Media Wall 1) showed the personalized collection of neutral pictures, e.g.

nature or sightseeing views containing no people or the user alone. The medium privacy version (Media Wall 2) showed the pictures containing the user and friends, but with no confusing content. Finally, the high privacy version (Media Wall 3) exposed the pictures with some compromising content: the user and friends in late party scenes, beach bikini scenes, etc.[4]

Each prototype provided five adaptive strategies aimed at privacy protection. Here, as private data are denoted: peers' names for Friend Finder 1, peers' names and pictures for Friend Finder 2, personal collection for Media Wall. The adaptive strategies were executing the following actions:

1. Do Noting: private data remained on the display.

2. Minimize: private data was shrunk in size, but remained on the display.

3. Mask: private data was occluded with solid blinders.

4. Remove Private Part: private data was completely removed from the screen. The neutral elements, such as uniform icons (Friend Finder) and working space (Media Wall) remained on the public screen.

5. Remove All: all private data was removed; the screen showed only the map.

The prototypes were presented in *PD-only* and in *PD-mobile* settings. The *PD-mobile* setting supported the strategies ***3-5*** where private data was not visible on the public screen: the private data migrated to the mobile screen, enabling the users to continue interaction.

All in all, ten prototypes were presented to each test participant: Friend Finder 1 and 2, in *PD-only* and *PD-mobile* setting, and Media Wall 1,2, and 3, also in *PD-only* and *PD-mobile* setting.

6.3.2 Experiment Procedure

Seventeen persons participated in the experiment, 7 female and 10 male, aged from 23 to 37 (average 29,3). Among them there were Italians, Russians, Chinese, and Germans, with the background of marketing, engineering, and multimedia research. All of them have experiences with social networks, such as Facebook, Studi-vz, XING, LinkedIn, myspace, InterNations; 12 persons have online photo collections.

At the beginning of the experiment the author introduced shortly the topic of adaptation on public displays, demonstrated possible adaptation strategies, and presented two applications, Friend Finder and Media Wall.

Then every participant was asked to imagine three individuals: a close friend, an acquaintance, and a stranger. The friend and the acquaintance should be thought of as real persons (the name was asked): the friend – a close trusted person and the acquaintance – a neutral familiar person, e.g. a colleague or a neighbor. The stranger was described by a uniform portrait: a male unfamiliar person, in his forties.

[4] The estimation of privacy levels was verified during the experiment; it matched the estimation of the participants.

In the main part of the experiment the participants were asked to evaluate the adaptation strategies for the prototypes, first Friend Finder and then Media Wall. Every prototype was shown first in *PD-only* setting, and then in *PD-mobile*. The order of privacy levels within prototypes was counterbalanced. For every prototype the author first demonstrated all five adaptation strategies. Then the author asked the participants to imagine the following scenario. The participant interacts with the public display alone and uploads the private data. Then a spectator suddenly approaches the display. The participant was asked which of the presented adaptation strategies would be preferred if the spectator was the friend, the acquaintance (named), or the stranger.

6.4 Results

The preferences of the participants were noted down as numbers from 1 to 5, referring to the strength of protection strategies (1 = ***Do Nothing***, 5 = ***Remove All***). The results were analyzed statistically, by comparing the preferences in various prototypes with pared t-test. Below the author reports on the results and gives comments.

6.4.1 Relationships Matter

As assumed, participants consistently chose stronger protection strategies for less close relationships. Significant differences were obtained through all the results for Friend Finder and Media Wall, in *PD-only* and *PD-mobile* settings (see Fig. 6.3).

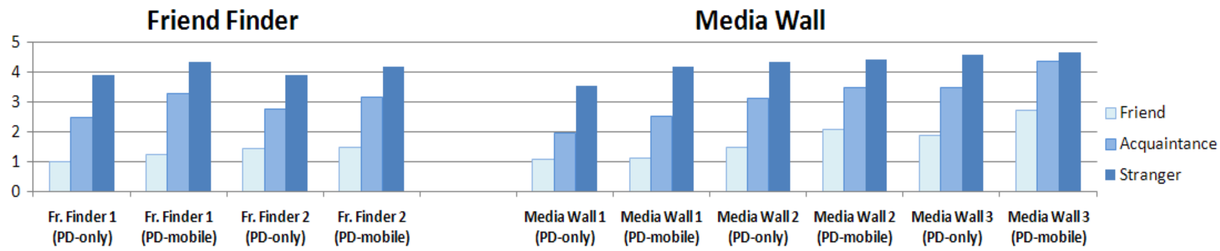


Figure 6.3. Preferences in protection strategies for different relationships with spectators.

Only for the stranger observing Media Wall 2 or Media Wall 3 in *PD-mobile* setting, the participants chose the strongest protection for both privacy levels.

6.4.2 Relationship Context in Friend Finder

Figure 6.4 summarizes the preferences in protection strategies for Friend Finder. The mean values are indicated above the graph bars.

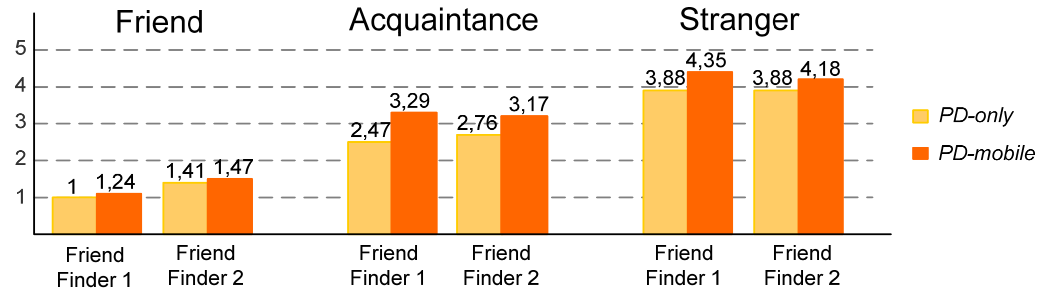


Figure 6.4. Preferred protection strategies for Friend Finder in *PD-only* and *PD-mobile* settings.

Spectator = Friend

Different privacy levels. The higher the privacy level of the content, the stronger protection strategy is chosen. Thus, the protection strategies for Friend Finder 2 were significantly higher than for Friend Finder 1 (p = 0.024). Friend Finder 2 discloses definitely more data on the social network: while just a name (Friend Finder 1) can stand for several persons, a picture and a name (Friend Finder 2) disambiguously points at a certain person. Often a spectator-friend shares some contacts of the user. The user might not be aware of private situation between the friends. Therefore, the user might prefer to hide the connections, in order not to confuse the common friends or the spectator: "Perhaps I have his [spectator's] ex-girlfriend in my network, and I have no idea what's their relationship now".

PD-only vs. PD-mobile Setting. The presence of a mobile device does not influence the choice of the protection strategy. If the users concern about the disclosure of their contacts, they choose a stronger protection in both settings.

Spectator = Acquaintance

Different privacy levels. No significant differences were found for privacy levels: similar strategies were chosen for Friend Finder 1 and Friend Finder 2.

PD-only vs. PD-mobile Setting. In *PD-mobile* setting the participants chose stronger protection than in *PD-only* setting. Friend Finder 1 (p = 0.0056) and Friend Finder 2 (p = 0.034). In the *PD-only* setting the users often sacrifice their privacy concerns for the sake of interaction comfort. Even under observation of an acquaintance, users choose less protective strategies which still enable them to proceed with interaction. The presence of a mobile device, however, eliminates the need to sacrifice the privacy; the mobile device enables further interaction and secures the private data. Therefore, a stronger protection is chosen on public display.

Spectator = Stranger

No significant differences were found for the stranger case. For both settings, *PD-only* and *PD-mobile*, independently on the privacy level, the preferences were spread between the stronger strategies ***3-5,*** which ensure invisibility of the private data.

6.4.3 Relationship Context in Media Wall

Figure 6.5 summarizes the preferences in protection strategies for Media Wall, showing the mean values above the graph bars.

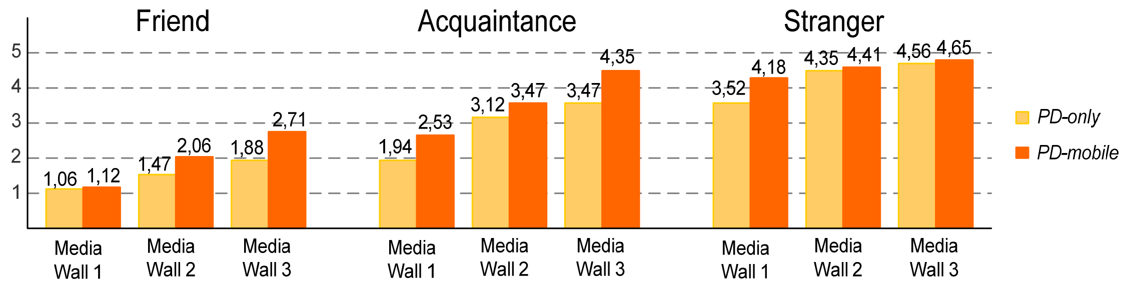


Figure 6.5. Preferred protection strategies for Media Wall in *PD-only* and *PD-mobile* settings.

Spectator = Friend

Different privacy levels. If a friend is observing the display, a strong protection is necessary only for highly private data. Thus preferences for Media Wall 3 were significantly higher than for Media Wall 1 (p = 0.0072) and for Media Wall 2 (p = 0.045). Highly private media often contain private information not only about the user, but also about their friends. Therefore, users prefer to hide the media in order not to confuse their friends who are even not aware of the possible disclosure.

PD-only vs. PD-mobile Setting. The protection preferences were significantly lower in *PD-only* setting than in *PD-mobile* setting, for medium (p = 0,038) and high (p = 0,015) privacy level. This result can be again explained by users' readiness to sacrifice their privacy concerns for the sake of interaction comfort: in *PD-only* setting users choose a weaker protection which does not impair the interaction. If a mobile device is available, the users continue the interaction on the mobile device and set a stronger protection on public display.

Spectator = Acquaintance

Different privacy levels. The higher the privacy level, the significantly stronger strategies were chosen, throughout all privacy levels.

PD-only vs. PD-mobile Setting. Low and high privacy levels require significantly stronger protection in *PD-mobile* setting than in *PD-only:* Media Wall 1 (p = 0,01) and Media Wall 3 (p=0,021). The protection for medium privacy level strongly depends on the role of the acquaintance. Users tend to decide once, if it is acceptable to show the content to the acquaintance and hold the decision for any setting.

Spectator = Stranger

Different privacy levels of content matter in *PD-only* setting. Medium and highly private data need significantly stronger protection than the low privacy level.

PD-only vs. PD-mobile Setting. The mobile device influences user decision only for low level privacy data: a stronger protection is chosen in PD-mobile (p = 0,043). For other privacy levels, the highest possible protection is chosen in either setting.

6.5 Discussion: Applying the Results

The results obtained in the experiment can be summarized as follows:

- **For a Friend-Spectator**, generally no protection is needed. The privacy concerns arise only if the display content can confuse the spectator-friend or compromise the persons involved in the content. In *PD-only* setting the users still keep the data opened, since hiding or removal is likely to impair the interaction process. In *PD-mobile* setting users tend to choose a higher protection: the mobile display serves as a safe depot for private data and enables further interaction with the content.

- **For an Acquaintance-Spectator**, a stronger protection is required. However, the preferences can be widely spread. Such distribution is caused by diverse roles of acquaintances. For instance, users may expose their holiday pictures to a neighbour, but prefer to hide them from a colleague. Having a mobile device available, the users tend to choose a stronger protection.

- **For a Stranger-Spectator**, the strongest protection is preferred. Since the users are not aware of intentions or interests of the stranger, they prefer to protect even the data of a low privacy level. The presence of a mobile device barely influences the protection preference: the users choose the strongest protection in either setting.

The obtained results can inform the design of a real-time relationship-based adaptation. The relationship information can be retrieved from the structure of user's social network (such as Facebook), from the intensity of chat and phone conversations. The identity of the spectator can be identified in real time by camera-based face recognition or by means of mobile phone ID. Additional context analyzed in the experiment can be also retrieved automatically. The setting context can be derived from availability of a mobile device. The privacy level can be extracted from the display content. For instance, if several faces are detected on a picture, the picture is automatically set to medium or high privacy.

Applying the results of the study to Friend Finder, a relationship-sensitive version was implemented. The network of friends was serving as a basis for extraction of friendship information. The network distinguished groups of friends and acquaintances; two not connected persons were considered strangers. By means of OpenCV library [159] adaptation was extended by face recognition algorithm. The algorithm was trained with a set of pictures of the social network members. In real time it was able to calculate the most probably match of the detected face and one of the network members. The name of the best matching network member and the respective probability were displayed on the camera picture.

The face recognition was embedded into the adaptation process as follows: When a person was watching his or her social network alone, the friends' portraits and names were shown (see Fig. 6.6 top). When a friend was approaching the screen, the social network remained exposed. When an acquaintance was approaching the screen, the portraits and names were minimized in size. When a stranger was close to the screen, the portraits and names were masked (see Fig. 6.6 bottom).

Important to notice that such adaptation can be used only if the face recognition delivers reliable results. In the current system, the face is considered as reliably recognized, if the classification probability is higher than 85%. If the probability is lower, the system considers the face as a stranger, and thus follows the most privacy-protective strategy: masking the faces. The user might be annoyed by the unnecessary protection of the data, when a friend's face is not recognized. However, the system thus assures that private information will never be exposed to a stranger.

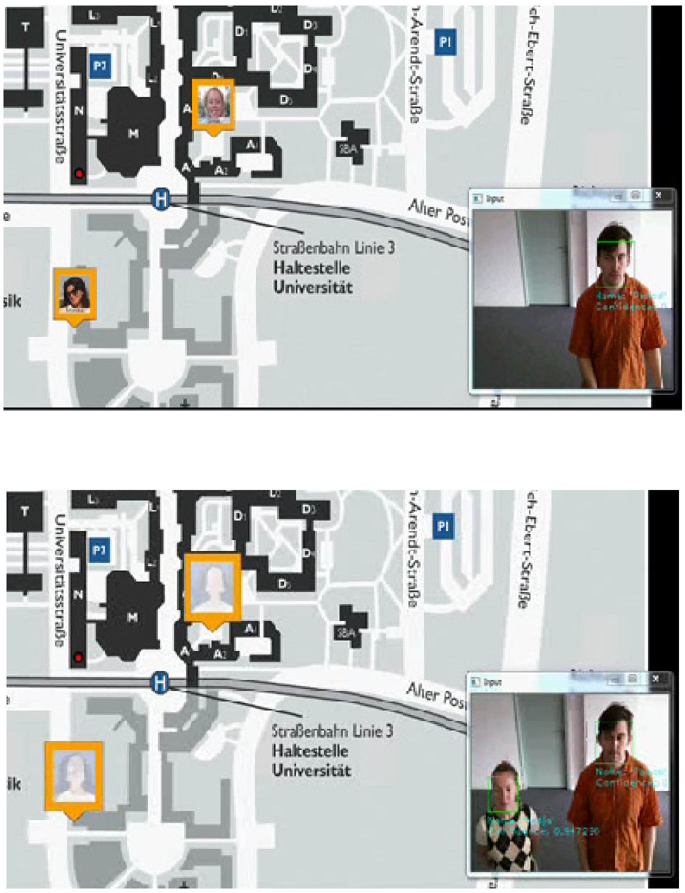


Figure 6.6. Adaptation of Friend Finder based on face recognition.

6.6 Summary

Relationship context should be taken into account when adapting public displays to a user's needs. By means of the literature review and the conducted experiment, the author showed that personal relationships with spectator significantly impact the user's preference in adaptation strategy. Using example applications, the author analyzed how this preference is influenced by privacy level of content and by the presence of an assisting mobile device. The reported results can guide the designers in creation of intelligent relationship-based adaptation mechanisms.

Important to notice that privacy concerns vary greatly among the users; they depend on the personality, traits of the character, personal experiences and can be summarized as *trust disposition* [130]. The experiment revealed that independently on nationality, gender, or age, participants showed some trust patterns, e.g. some of them concerned more about privacy in social networks, others – about private pictures. Therefore, the definition of "universally applicable" privacy levels still remains a challenging task.

The preferences found in the experiment can serve as recommendations for adaptation design for the diverse contextual settings. However, designers should always provide the users with leverages to override the automatic adaptation, so that the users feel the ultimate control over the system behaviour.

Chapter 7

Automatic Decision Making for Adaptation on Public Displays

Ubiquitous display environments require a high degree of flexibility due to the changing social context. Modern displays may intelligently adapt, for instance, to the profiles of spectators, their gender or mutual proximity. However, the adaptation decisions should maintain the user's trust in the system. A wrong decision can negatively influence the user's acceptance of a system, cause frustration and as a result, make users abandon the system. In order to support user trust in such environments, the system should be able to evaluate the consequences of its actions. In this chapter the author presents a trust-based mechanism for automatic decision making. The mechanism assesses the impact of system actions on user trust in critical contexts and enables to find the best fitting adaptation decision automatically.

7.1 Motivation

Recent years have brought about a large variety of interactive displays that are installed in many public places. Apart from simply providing information (e.g. news or weather), public displays enable individuals to personalize, exchange, and edit the exposed data. Although such adaptation ability offers convenient service to the spectators, the exposure of personalized information in a public place inevitably raises privacy issues. Apart from the privacy issues, the adaptation risks to mismatch the expectations of the spectators. Imagine a display which automatically distributes space for convenience of its spectators. A newly coming spectator expects that a free space slot will appear next to his position. When the new person is approaching the display, the system must take a decision: Provide a space slot? Where exactly to provide it? Protect personal data of existing spectators? Or ignore the person, since maybe he is just passing by?

If the system decides to execute an adaptive action, it might disturb existing spectators: their workflow may be interrupted. If the system does not execute any adaptation, the new comer spectator may be upset; his trust into the system will be diminished.

The example illustrates that a system needs to carefully balance the benefits and drawbacks of its actions in order not to risk that a user loses trust and abandons the system.

Another source of risk in making decision on adaptation action is the uncertainty of social context. Although modern sensor technologies provide advanced means for context recognition, the acquisition of precise context data still remains a challenge. Indeed, face recognition may be not robust to non-frontal faces and speech recognition suffers from the unpredictable noise of a public place. Therefore, the tasks such as gender or age recognition can be estimated only with a certain probability. As a result, the display has to take a decision on adaptation based on an incomplete picture of the surrounding context, or in other words, under uncertainty of social context.

In this chapter the author presents a mechanism for automatic decision making based on a trust-based Bayesian Network (BN). Relying on empirical data, the author initializes the network and demonstrates how it can be used to generate adaptation decisions under changing social context.

7.2 Making Decisions on Adaptation of Public Displays

The adaptation strategy of most public displays is usually pre-defined. The defined mechanism relies on the recognized incoming context; the system matches the context to one of the pre-defined adaptation actions. In other words, in order to generate a decision on an adaptation action, the display needs to find an exact match between the sensed content and the predefined adaptation action. Such systems, however, are not flexible to handle situations with deviating or uncertain context. The deviating or uncertain context is usually handled by default adaptation actions or idle states. Such actions, however, often do not match the expectations of the spectators.

Alt and colleagues presented an intelligent mechanism to choose the content for a public display based on the interests of the surrounding spectators [6]. The adaptation of content is based on the spectators' profiles transmitted to the display via the spectators' mobile devices. The authors, however, do not address situations where a part of the spectators is not recognized. Although these unrecognized people also observe the content, their interests are not considered when taking a decision on the new content.

Exeler and colleagues presented an ambient display aimed to reflect audience emotions [54]. The emotions were derived from a camera-based recognition of facial expressions. The display showed a face performing the same emotion as the one recognized by the camera. The system, thus, aimed to entertain and show participation in the audience's mood. However, it did not provide any solution for poorly recognized or undefined emotions. In case the system could not recognize the audience's emotions, the face on the display remained neutral. In these cases the goal of the installation could be threatened: although the audience is highly emotional, the system fails to recognize it, and the face on the display does not react at all. One can imagine the frustration and reduced user acceptance which such a system behavior could lead to.

Müller and colleagues propose an automatic decision making mechanism to the content selection on a public advertisement display [146]. The user or composition of users watching the advertisement was identified by means of face recognition or mobile sensors. In order to generate the decision on the suitable content, the system considered the statistics on general purchases of the advertised item and the purchase history of every user. Thus, the system aimed to increase the sales and avoid showing irrelevant information to the users. The approach presented an intelligent way to treat uncertain social context: if the user profile could not be retrieved, the decision on the adaptation is based on the statistics of the product sales. The mechanism, however, was based on a default state – the advertisement of a best seller product – and not on the learnt behavior of the spectators.

To summarize, the existing public displays usually rely on pre-defined adaptation actions which are mapped to the recognized social context. Unknown or uncertainly recognized context is usually mapped to default adaptation actions or idle state, i.e. no adaptation at all.

In order to find out how to design adaptation when the context is unknown or uncertain, the author took a look at research on adaptation in other domains.

Horvitz and colleagues proposed a mechanism to detect user attention under uncertainty of context [87]. The authors utilized multiple sensors to learn and predict the user's attention state, such as calendar events, email texts, microphones, and cameras [88]. The mechanism aims to determine whether an interruption (such as an alert or pop-up message) will be appreciated by the user or should better be postponed. In order to draw conclusions about the user's immediate attention resources, the system utilizes a Bayesian Network [86]. The network is initialized and trained by empirically collected data, and thus is tailored to each user individually. In case of uncertain context, such as poorly recognized speech, the network generates a solution on interruption possibility based on the remaining available context and the learnt behavior of the user. This way, the approach is able to cope with the problem of decision making under uncertain context.

Inspired by this approach, the author designed a decision-making mechanism for adaptive public displays based in Bayesian Network (BN). Similarly to the work of Horvitz and colleagues, the network was aimed to predict an appropriate adaptation reaction even when only partial knowledge about the context is available.

7.3 Building Bayesian Network

A critical factor which affects the users' acceptance of an adaptive display and which impacts the users' decision to use the system ever again is the users' trust towards the system [181]. Therefore, the theory of user trust [181] was taken as a basis for the construction of the BN.

The first step on the way to the construction of the BN was to identify trust dimensions, i.e. critical design factors that contribute to the formation of trust [110]. Having the trust dimensions, it was possible to model the network defining the nodes and their dependencies.

7.3.1 Identifying Trust Dimensions

In the research on trust from the areas of sociology, psychology, economics, and computer science, there is a consensus that trust depends on a variety of trust dimensions. The set of dimensions is, however, not fixed.

In the area of e-commerce, Grandison and Sloman [72] or Kini and Choobineh [104] use trust dimensions reliability, dependability, honesty, truthfulness, security, competence, and timeliness. In the area of sociology, researchers [206] employ the trust dimensions benevolence, reliability, competence, honesty, willing, vulnerability, and openness.

Researchers working on adaptive user interfaces consider transparency, controllability, privacy, comfort of use, and reliability as major facets of trust.

Transparency and Controllability [69]. Research on trustworthy design indicates transparency and controllability as the main aspects supporting user trust. Glass and colleagues emphasized the importance of transparency and control in the design of trustable agents [2]. Graham and Cheverst studied interaction paradigms that maintain trust in mobile guides [42]. The authors identified that the lack of transparency and control potentially diminishes user trust. Cheverst and colleagues designed a system that dynamically adjusts to a learnt user model [54]. The authors emphasize the importance of sufficient transparency and comprehensibility of the system and the need to control the existing user model. Bellotti and Edwards [18] as well as Lim and Dey [125] claim that intelligibility significantly improves user trust in context-aware systems. If private data is involved in the adaptation, transparency and control gain even a greater importance. Langheinrich claims that ubiquitous systems should explicitly inform users of aspects that relate to their privacy [118, 119]. The users should be empowered to cancel unauthorized actions.

Privacy [2, 120]. Research works on public displays often point at privacy as a potential cause of interaction and acceptance issues. Röcker and colleagues studied how people use public displays which expose personal information [181]. The authors found out that users appreciate the use of displays in public settings, profiting from the large screen estate. However, they are worried about their privacy protection.

Comfort of Use. Although transparency and control evidently support user trust, their excess can negatively impact interaction comfort. Frequent confirmation requests, choice options, and explanations of the adaptive behavior may hinder a fluent interaction process and, as consequence, make the user abandon the system.

Reliability [181] is rather a basic dimension. The system is required to run in a stable manner.

The set of trust dimensions derived from the literature was verified and enhanced based on interviews with 20 students of computer science. The students were asked to indicate the factors of user interfaces that they felt contributed to their assessment of trustworthiness. The most frequent mentions felt into the following categories: comfort of use ("should be easy to handle"), transparency ("I need to understand what is going on"), controllability ("want to use a program without automated updates"), privacy ("should not ask for private information"), reliability ("should run in a stable manner"), security ("should safely transfer data"), credibility ("recommendation of friends") and seriousness ("professional appearance").

Trust depends on experience and is subject to change over time. Lumsden [130] distinguishes between immediate trust dimensions and interaction-based trust dimensions. Immediate trust dimensions, such as seriousness, come into effect as soon as a user gets in touch with a software system while interaction-based trust dimensions, such as transparency of system behavior, influence the users' experience of trust during an interaction. There is a consensus that trust is highly subjective. A person who is generally confiding is also more likely to trust a software program. Furthermore, users respond individually to one and same event. While some users might find it critical if software asks for personal information, others might not care.

The interviews gave an overview of the factors that influence user's trust in an adaptive system. However, the factors do not provide any concrete information on their relative importance. To gain more understanding on the relation of the factors, an empirical study was conducted [110]. The study revealed that there are statistically significant positive correlations between trust and the identified factors. The better the ratings for the trust dimensions, the better were also the ratings for trust. Furthermore, the author found that a missing feeling of trust was accompanied by negative emotions, such as irritation, uneasiness and insecurity.

7.3.2 Modeling the Bayesian Network

To model trust, a decision-theoretic approach was applied. The basic idea was to define factors that have an influence on the user's feeling of trust and to investigate how these factors can be influenced by particular system actions. The users' feeling of trust was modeled by means of Dynamic Bayesian Networks [187].

The structure of a Bayesian Network is a directed, acyclic graph (DAG) in which the nodes represent random variables while the links or arrows connecting nodes describe the direct influence in terms of conditional probabilities. Bayesian Networks enable to model the influence of different trust dimensions on the user's trust in a rather intuitive manner. For example, it is rather straightforward to model that reduced transparency leads to a decrease of user trust. The exact probabilities are usually difficult to determine. However, the conditional probabilities can also be (partially) derived from the user data (see [17] for an example of an experimental collection of this data).

As a test bed for the research, two applications were employed: Friend Finder and Media Wall. The applications have been developed as part of a university-wide displays management system. The two applications run on public displays located in public rooms at Augsburg University. They can be operated and assisted by mobile phones. The detailed description of the applications can be found in the Sections 5.1 and 6.3.

The first application, Friend Finder, is an interactive campus map that shows the current location and status of the user's friends. Friend Finder also supports a routing function, showing a detailed path to a selected friend (see Fig. 7.1).

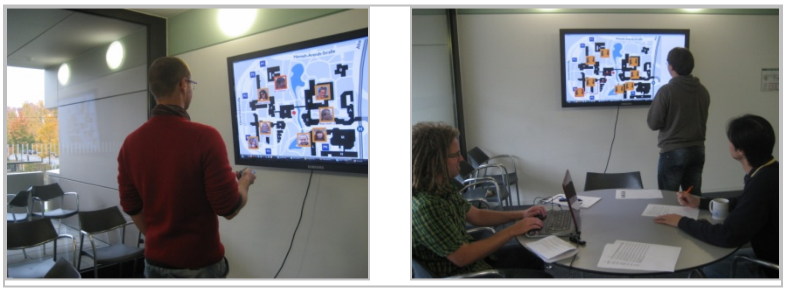


Figure 7.1. Friend Finder application, in a public setting.

The second application, Media Wall, fosters media exchange between the students. Users can rank the media items, upload new items, and view their favorite ones (see Fig.7.2).

Both applications require sophisticated mechanisms to adapt to various trust-critical events. Since Friend Finder may disclose private information about user's social network, it should be able to intelligently adapt to the surrounding social context in order to avoid possible privacy threats.

For example, the user might not feel comfortable to view personal data on the public display in the presence of the passer-by behind her. Ranking the media on Media Wall again may threaten the user's privacy in case of observation. Several users may interact with Friend Finder simultaneously, rendering their networks on the same campus map. Therefore, the system should be able to accommodate the data and interaction coming from multiple users.

Figure 7.2. Media Wall application, in a public setting.

Figure 7.3 illustrates a part of the Bayesian Network; the network was used for modeling trust in our two applications: Friend Finder and Media Wall.

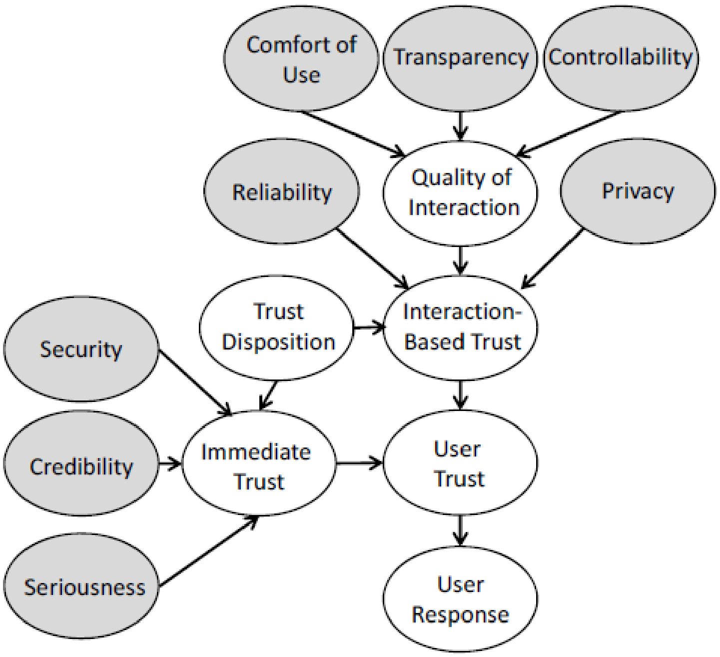


Figure 7.3. Bayesian Network with trust dimensions (trust dimensions in gray)

The left part of the network represents the factors influencing the establishment of *Initial Trust* that arises when a user gets a first impression of a system. *Initial Trust* consists of the trust dimensions *Security, Seriousness* and *Credibility. Security*, for example, could be conveyed by the use of certificates. A system's *Seriousness* is reflected, for example, by its look-and-feel. *Credibility* could be supported by additional information, such as a company profile. Because these factors change rarely and for the sake of simplicity, the author assumes that the *Initial Trust* and its dimensions do not change over time. However, the establishment of *Initial Trust* is also influenced by the user's *Trust Disposition* including the user's *Competence* and general *Confidence* into technical systems. The user's *Trust Disposition* also influences the development of *Interaction-Based Trust* that is modeled in the right part of the BN.

The determinants of *Interaction-Based Trust* can be subdivided into the five trust dimensions the author identified as relevant to the field of adaptive public displays: *Privacy*, *Reliability* as well as *Transparency*, *Controllability*, and *Comfort of Use* that characterize the *Quality of Interaction.*

Dynamic Bayesian Networks allows, in addition, modeling the dependencies between the current states of variables and earlier states of variables. In particular, it enables to represent how the user's current level of trust is influenced by earlier levels of trust. To keep things simple, let's only consider the user's level of trust at time $t_{i-1}$when determining the user's level of trust at time t_i (see the extension to a Dynamic Bayesian Network in Fig. 7.4).

To use the Bayesian Network formalism for decision-making, it has to be extended to an influence diagram by adding a decision node and a utility node. The decision node represents all system actions that the system can perform while the utility node encodes the utilities of all possible outcomes.

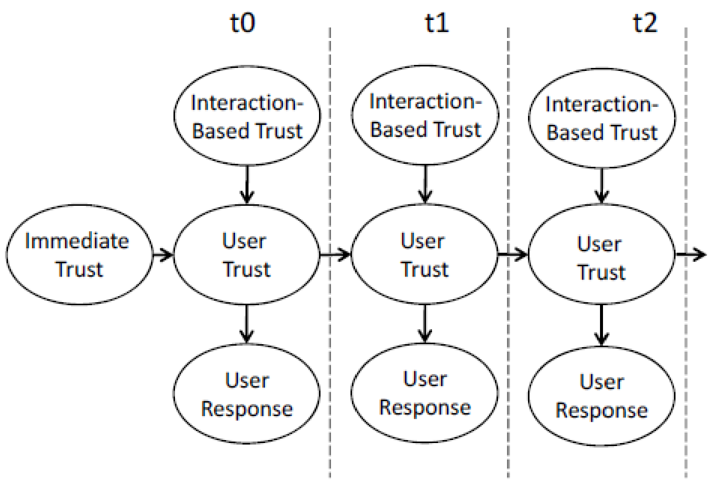


Figure 7.4. Dynamic Bayesian Network

To take a decision, the system evaluates the utility of all possible options in terms of user trust and chooses the action with the highest utility. In Fig. 7.5, a small portion of the influence diagram is shown. The figure illustrates the basic idea by means of one trust dimension, namely privacy.

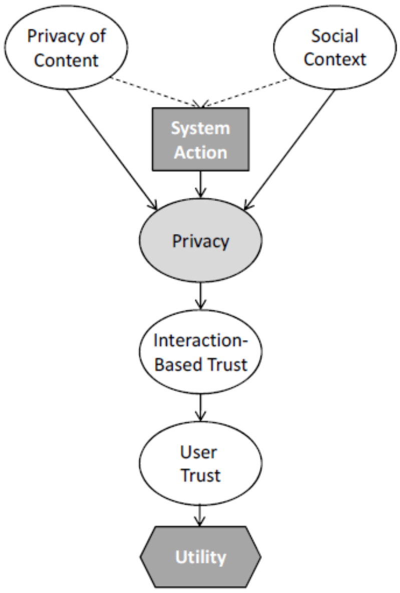


Figure 7.5. Small portion of an influence diagram.

Privacy is handled as a hidden variable with three discrete values low, medium and high. That is its value cannot be directly observed, but has to be inferred from observables variables, such as *Privacy of Content* and *Social Context*. For example, the likelihood that the variable *Privacy* has the value *Low* would be high if *Privacy of Content* has the value *Private* and *Social Context* has the Value *People Approaching*. These dependencies are indicated by the arrows going from *Social Context* and *Privacy of Content* to *Privacy*. Associated with the decision node *System Action* is a table which describes the system's decision policy for each combination of the variables Social Context and Privacy of Content. The arrow going from System Action to Privacy represents the impact a particular system action, for example, the masking of private information, has on privacy.

Important to mention, Dynamic Bayesian Network was used to illustrate the concept of the decision making mechanism. The data collection and the experiment were conducted for a non-dynamic network. The non-dynamic network represents a fragment of the Dynamic Bayesian Network (see Fig. 7.4) for a certain time moment. The goal of the empirical data collection and the experiment was to demonstrate how the network can be used to generate adaptation decisions based on user trust. Therefore, a fragment of the Dynamic Bayesian Network is sufficient to illustrate the decision-making mechanism. The non-dynamic fragment can be later integrated into the chain of the Bayesian Networks, in order to arrange a Dynamic Bayesian Network. Such a network will be able to take adaptation decisions based on the earlier values of user trust.

7.3.3 Setting Probabilities

The following section illustrates how to set the probabilities for concrete applications in the described BN. The process of building up conditional tables can be found in [17], where the authors present an experiment on data collection for modeling user trust. In the experiment the users had to rate trust and trust factors for prototypes they got confronted with.

In this section the author will focus on setting the probabilities for the decision-making process. One possibility is to learn influence diagrams based on collected user data. Another option is to set up the influence diagrams based on informed guesses. For example, probabilities that represent the influence of actions on trust factors referring to the quality of interaction may be assessed by considering usability guidelines. Since the acquisition of data is rather time-consuming and can in most cases not directly be transferred from one application to the other, the second approach will be followed.

The analysis of the impact of various system reactions to typical trust-critical situations on relevant trust factors is illustrated by the example of Friend Finder and Media Wall.

Let us assume that the user is viewing private data on the public screen as other users pass by. Such a situation may occur in Friend Finder when users load a map of the university campus with friends on a public screen. The locations and pictures of friends are considered as private information, not supposed to be observed by any one. In the influence diagram shown in Fig. 7.5, this situation is described by the values of the variables *Social Context* and *Privacy of Content*.

Four possible responses to the described situation are listed below. Basically, the system has to decide whether it should:

a) adapt the display to the changed social context by masking the data on public display and moving all private data to a mobile display (Option 1),

b) adapt the display to the changed social context by moving private data partially to a private device (Option 2),

c) ask the user for confirmation first (Option 3),

d) not undertake any action (Option 4).

In the influence diagram shown in Fig. 7.5, the available options are represented by the decision node *System Action*.

Option d) bears the risk that the user might expect the system to protect her privacy and is upset if no appropriate actions are taken. Options a) and b) have the limitations that they cause an interruption of the user's work flow and might give her the feeling that she has the system no longer under control. The drawback of Option c) is that there might not be enough time for the user to confirm the adaptations proposed by the system.

Compared to Option a), Option b) has the advantage that the user is still able to execute desired actions. Thus, for instance, on Friend Finder the user can select a desired friend and get a route to his or her destination. In this case, the user still profits from the large real estate of the public screen, while preserving personal information. Furthermore, the Option b) allows several users to interact at the same time (see Fig. 7.6).

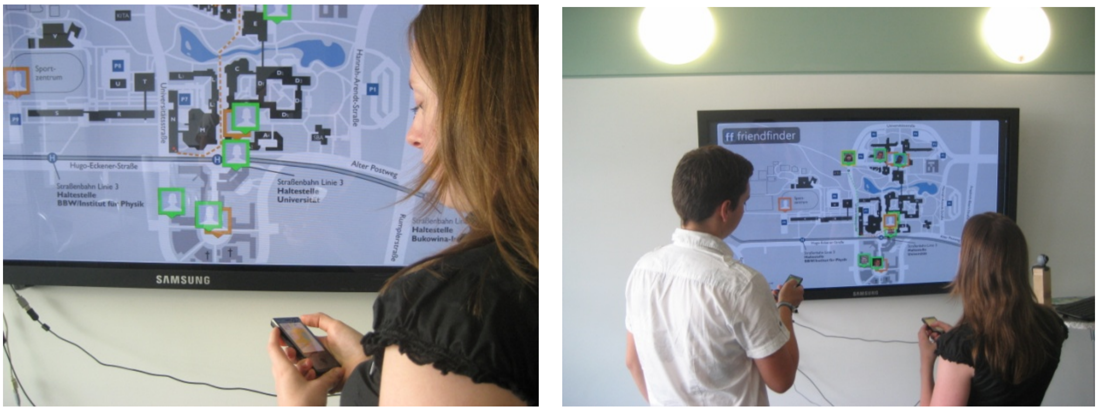


Figure 7.6. Multiple users interact with the public screen using mobile displays.

Table 7.1 summarizes the analysis above. It lists the consequences of system actions on relevant trust dimension. In particular, the author indicates how likely it is that the system actions will change the ratings of the trust dimensions if private context is displayed in the presence of others. Based on this table, conditional probabilities may be derived that represent the dependencies between system actions and trust dimensions. For example, the likelihood that privacy is low is high if private content is displayed in the presence of other people and the system takes no action while it is low if the system masks private content. Please note that the table refers to a particular situation that is described by values of the variables *Social Context* and *Privacy of Content*. For other situations, additional tables describing the impact of possible system actions on trust dimensions need to be built up.

	Transpa-rency	**Controll-ability**	**Comfort of Use**	**Privacy**	**Reliability**
No Adaptation	0	0	0	- -	- -
Complete Migration	- -	- -	+	+	0
Partial Migration	-	-	-	+	0
Ask for Confirmation	+ +	+	- -	+ +	0

-: Somewhat likely to decrease, - -: likely to decrease, 0: not likely to change,

+: Somewhat likely to increase, + +: likely to increase

Table 7.1. Relationship between system actions and trust dimensions

In order to be able to generate decisions, the BN needed to be initialized with empirical data. The data was collected through experiments conducted with potential users. The users were confronted with scenarios illustrating different contextual combinations and possible adaptive reactions of the displays that differed in the degree of transparency, user control, privacy and comfort of use. To discover which of the system reactions succeeded in maintaining the users' trust and which did not, the users had to reflect on their perception of the display reactions and had to give insights into their feelings of trust and the related trust dimensions. The estimations served as a quantitative input for the initialization of the BN.

The collection of empirical data was arranged in two steps: First, the author conducted a *web-based study* targeting as many users as possible. This study presented a collection of applications demonstrating various content types typical for modern public displays: social networks, sharing of pictures and videos, maps and travel planning, and shopping items. These content types can be frequently found in the real life projects [132, 165] as well as in research works [6, 42]. The applications showcased different situations where an adaptation was necessary: space conflicts, privacy protection, and migration of data from public to mobile displays. The reason to involve several applications was based on our objective of verifying that the proposed approach works equally (or comparably) well for different kinds of content, different sources of social context, and different adaptation scenarios. If the approach indeed delivered robust results, its applicability could be generalized to a wide range of adaptive applications.

Since a web-based study might not transmit the experience of a real interaction and thus affect the reflections of the users, the author additionally conducted controlled *live experiments*. The experiments replicated two scenarios presented in the web-based study and involved two different applications. The live experiments were designed identically to the web-based study, but involved real user interactions.

All in all, the web-based study was aimed to gather as much data as possible, involving online users. The live study was aimed to complement the web-based study, supporting the results collected online by the evaluations of users during a real interaction. Below the author describes both studies in detail and present the obtained results.

7.3.4 Web-based Study

The web-based study was aimed at capturing user perception of display reactions in scenarios with changing social context. The following adaptation scenarios were considered:

- Privacy issues: Adaptation aimed at privacy protection.

- Space conflict: Adaptation aimed at resolution of the space conflict.

The listed scenarios represent typical adaptation triggers which call for adaptation on public displays. Adaptation can involve only the public display, or also involve other available displays, such as mobile devices or tablet PCs. The scenarios took into account these different possibilities. The scenarios were illustrated with short videos, featuring four public display prototypes.

The first and second prototypes were Friend Finder (FF) and Media Wall (MW) described in the above section.

The third prototype, Travel Planner (TP) aimed to support students who plan their low-budget trips around Europe. By selecting a city on the large display map, the users activated a pop-up which showed them general information on the city and the estimated cost of the visit (see Fig. 7.7). Apart from this neutral information, the application could show private budget-aware data. In the budget-aware mode, the pop-up information was linked directly to the user's budget: it showed whether the selected city fitted into the user's limited budget and indicated the remaining amount of the user's money. The application supported two presentation techniques: the pop-ups could be shown either on the public display or on the user's mobile. Such a choice helped users protect their private data, depending on the mode. Travel Planner could be used by several people simultaneously.

Figure 7.7. The prototype of Travel Planner.

The fourth prototype, Shopping Mall Display (SM), aimed to inform customers of a shopping mall about the products they were looking for (see Fig. 7.8). Several customers could use the display in parallel; each of them would get a segment of the screen where the selected item was shown. When a new user approached the display, the screen automatically tried to provide him or her with a new space segment. The display could also offer its space, if a passing-by user was browsing for a product on a mobile phone or a tablet. When the user approached the display, the item from the mobile or table screen could be transferred to the large screen. Of course, this function might not always be appreciated, for instance, if the user was browsing for a private item.

For each application, several short videos were recorded. The videos illustrated different adaptation scenarios. Each video demonstrated a scenario in which the social context changed and the display reacted to it accordingly. For example, one of the scenarios for Friend Finder showed the user coming closer to the display in a public area. The display recognized the proximity of the user and showed the user's social network automatically. Another video illustrated the same scenario, but an alternative display reaction: instead of placing the data automatically, the display presented the user with various options to adapt to the new situation via the user's mobile phone.

Figure 7.8. The prototype of Shopping Mall Display.

Table 7.2 summarizes the recorded scenarios, the context which was represented by different settings of contextual variables, and the possible display reactions. The applications illustrating the scenarios are indicated by the capital letters in the Scenario column.

Scenario	Display Reaction	Context
1. User comes closer to the display. Adaptation trigger: privacy issues. *FF, MW, SM*	a) Show user data automatically b) Ask user for decision via mobile device c) Do nothing	*Privacy of data* - private / neutral *Spectators present* - yes / no
2. User leaves the display. Adaptation trigger: privacy issues. *FF, MW*	a) Remove user data automatically b) Ask for confirmation to remove, via mobile device c) Do nothing. Data remains on the display.	*Privacy of data* - private / neutral *Spectators present* - yes / no
3. User interacts with the display. **Another user approaches the display.** Adaptation trigger: space conflict. *SM*	a) Provide space for the new user, by shrinking the data of existing user. b) Provide space for the new user, by moving some data of existing user to his / her tablet device. c) Do nothing. The new user has to wait.	*Tablet available* - yes / no *Gender of new user* - female / male

4. User interacts with the display alone. **Suddenly another person approaches the display.** Adaptation trigger: - privacy issues. *FF, MW*	a) Hide private data b) Mask private data c) Do nothing. Private data remains on the display. d) Offer the choice between options a), b), and c) via mobile phone.	*Privacy of data* - private / neutral *Relationship with spectator* - friend - acquaintance - stranger
5. User interacts with the display alone. Adaptation trigger: privacy issues, space conflict. *TP*	a) Show content on public display b) Show content on mobile device	*Privacy of data* - private / neutral *Another user present* - yes / no

Table 7.2. Scenarios illustrated by videos: possible display reactions in different contextual combinations.

All in all, 68 short videos were recorded. For each scenario several videos were created, demonstrating all possible combinations of the display reactions and context. Thus, 12 videos were recorded for Scenario 1, 12 for Scenario 2, 12 for Scenario 3, 24 for Scenario 4, and 8 for Scenario 5. In order to reduce the time of the survey completion to about 10 minutes, the author grouped the videos into six surveys. Each survey contained about 8-12 videos. After an introductory page, the surveys provided a description of the used applications. Then, the user was confronted with the first scenario. The scenario was followed by a video, which illustrated the first display reaction to the shown situation of the context change. After each video the user had to fill in a questionnaire. The questions aimed at capturing the user's perception of the shown display reaction in terms of transparency, controllability, comfort of use, privacy, reliability, and trust. The questions represented statements which had to be ranked on a 5-point Likert scale:

Q1: "I understood why the system was reacting in this way."

Q2: "I had control over the system"

Q3: "I found the system comfortable to use"

Q4: "The system protected my privacy in an appropriate way"

Q5: "I found the system reliable"

Q6: "I found the system to be trustworthy"

After presenting all possible display reactions for a particular scenario, the users were asked to rank their preferences for it. The preferences also had to be estimated as statements to be assessed in 5-Likert scale. The statements emphasized the context of the given scenario, such as the presence of others or the privacy of data. For instance, a statement for the scenario of Friend Finder where the user was coming closer to the display looked like this:

"When I am watching my social network alone and a stranger approaches the display..."

P1: I prefer to hide my data

P2: I prefer to mask my data

P3: I prefer no reaction from the display

P4: I prefer to be asked by my mobile phone

Questions Q1-Q5 were aimed to collect empirical data to initialize the BN. Question Q6 was needed for validation of the network: to compare whether the decisions generated by the network match the display reaction with the highest user trust. Therefore, questions P1-P4 aimed to give us an insight how subjective user preferences match users' feeling of trust. In particular, the author wanted to find out whether the display reactions explicitly preferred by the users also get the highest trust ratings. Moreover, the author wanted to see whether user preferences match the decisions generated by the BN.

All in all, evaluations of 92 online users were collected. The online users were recruited by the authors of the experiment. The link with the web-based study was sent to them by email. Moreover, the link was published on the social network site of the department offering the visitors to participate in the study.

The participants of the online study offered to give their demographic data. This information was, however, not mandatory. The 73 users that provided demographic data included 35 women and 57 men. They were aged between 23 and 62 years, with an average age of 35.6 years.

The scenarios were distributed by the online users in the following way: 17 users evaluated privacy-critical scenarios with Friend Finder (Scenario 1 and 2), 13 users evaluated privacy-critical scenarios with Media Wall (Scenario 1 and 2), 18 users evaluated privacy-critical scenarios with relationship context with Friend Finder (Scenario 4), 12 users evaluated privacy-critical scenarios with relationship context with Media Wall (Scenario 4), 16 users evaluated space conflict and privacy-critical scenarios with Shopping Mall display (Scenario 1 and 3), 16 users evaluated space conflict and privacy-critical scenarios with Travel Planner (Scenario 5).

Before using the data collected in these online studies for the initialization of the BN, the author investigated whether the results of the online study were in line with the perception of users actually interacting with an adaptive system.

7.3.5 Live Experiments

Since online studies might not transmit the experience of a real interaction and thus affect the reflections of the users, the author verified the results of the online study by means of controlled live experiments. For the live experiments two scenarios from the online studies were picked randomly. The prototypes used for the experiment were Friend Finder and Travel Planner. The scenarios reproduced in the experiment included the context variables privacy of data and presence of other user/spectator. The experiment of Friend Finder was playing out the Scenario 1 and 2, when being alone in the public area or when other people were present. The experiment with Travel Planner reproduced the Scenario 5, with neutral or budget-aware (private) data, when being alone or when other user was present.

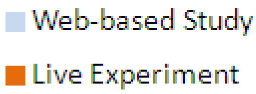
The experiments were conducted individually, in a real public area, in front of a large display. Both prototypes were tested as between group test, meaning that every user evaluated only one application, either Friend Finder or Travel Planner.

Analogue to the web-based study, in each experiment the participants first received a short description of the presented prototype. Then they were introduced to the scenario, emphasizing the given context, such as standing alone in front of the display, or being observed by other people. For each scenario the users were confronted with different display reactions, identical to those presented in the online study. After experiencing each display reaction, the participants were asked to fill in a short questionnaire where they estimated the display reaction in terms of transparency, controllability, comfort of use, privacy, reliability, and trust. The questionnaires were formulated identically to those from the online study.

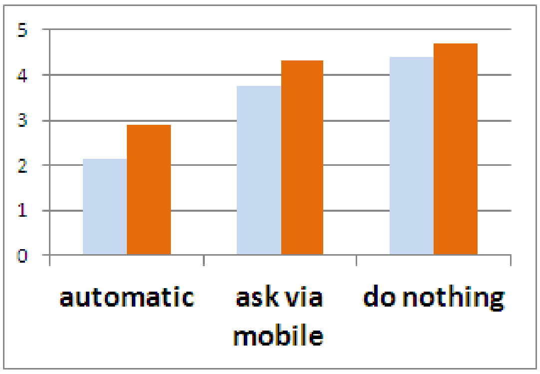
After the users had experienced all display reactions in the given context, they were asked to estimate their preferences. The questions on preferences were identical to the ones presented in the web-based study. Thus, the procedure of the live experiments completely replicated the procedure of web-based study.

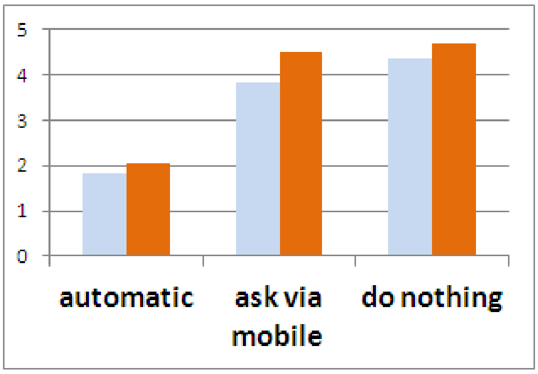
Altogether, 36 people took part in the live experiments (Travel Planner: 20, Friend Finder: 16). Among them there were 16 female and 20 male persons, aged from 20 to 36 (mean 28.3).

The results of the live experiment generally matched the results obtained in the web-based study. Both experiments yielded similar distributions of user rankings of transparency, controllability, comfort of use, privacy and reliability. Moreover, the author got similar distributions of trust (see Fig. 7.9) and user preferences. The Figure 7.9 illustrates the trust estimations in web-based and live studies, by the examples of Scenario 1, 2, and 5.



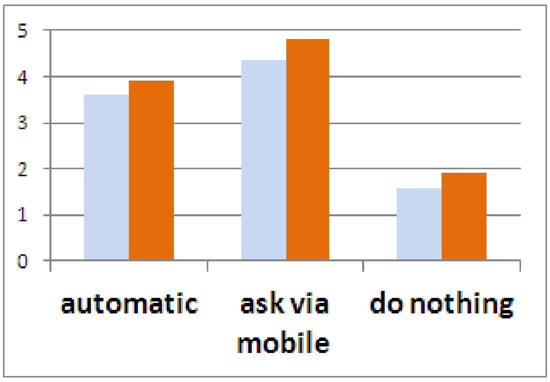
Trust: Friend Finder (Scenario 1)

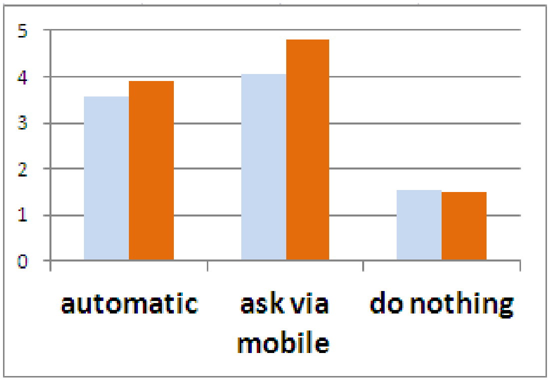




User alone **Spectator present**

Trust: Friend Finder (Scenario 2)

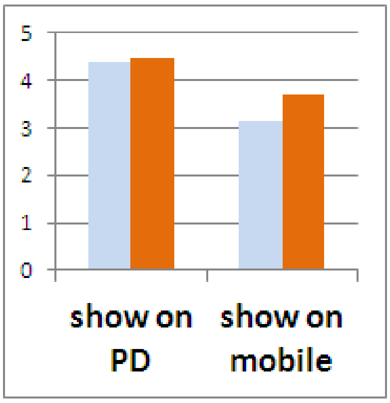




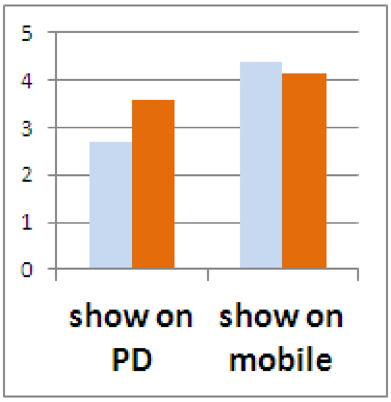
User alone **Spectator present**

Figure 7.9. Distributions of user rankings for trust in live and web-based studies.

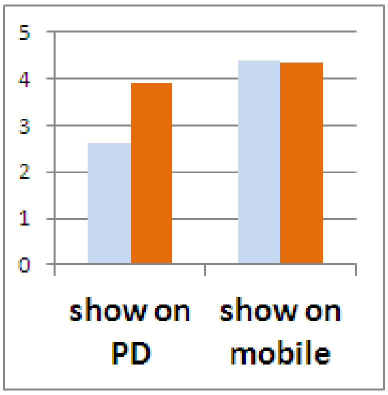
Trust: Travel Planner (Scenario 5)



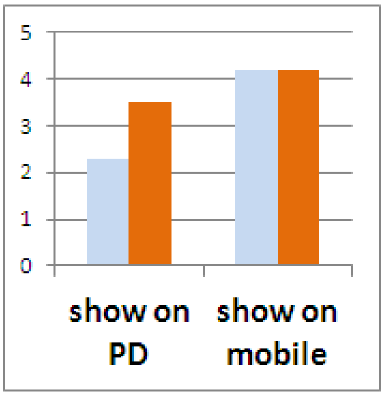
User alone, neutral data



Spectators present, neutral data



User alone, private data



Spectators present, private data

Figure 7.9 (continued). Distributions of user rankings for trust in live and web-based studies.

Overall, the trust ratings in the web-based and in the live experiments show a similar trend. The participants gave slightly higher trust ratings in the live condition than in the online condition. However, the differences were not significant. Apparently, the fact that the participants had the chance to interact with the system had influenced their ratings positively.

The important result was to see that apart from a few exceptions the ranking of system reactions in the online experiments was in line with that obtained in the live experiments. Independently of whether users had to evaluate the online or the live setting, the participants gave their preferences and the highest trust ratings to the same system reactions. The comparison of the ratings given to different system reactions was conducted by means of the two-tailed paired t-test.

In the case of Friend Finder, Scenario 1, the system reaction "Do nothing" got significantly higher rankings than the reaction "Ask via mobile device" (means 4,41 (STD=0,79) and 3,76 (STD=0,83), $t(18) = 2{,}32$, $p = 0{,}02$). The system reaction "Do nothing" also got significantly higher rankings than the reaction “Show automatically” (means 4,41 (STD=0,79) and 2,11 (STD=1,05), $t(18) = 5,06$, $p < 0,001$). The result illustrates the scenario with private data, user alone. A similar result was obtained for the scenario with private data, spectators present. Here, the system reaction “Do nothing” again was rated higher than the reaction “Ask via mobile device”, though not significantly higher (means 4,35 (STD = 0,86) and 3,82 (STD = 0,88), $t(18) = 1,77$, $p = 0,086$). The reaction “Do nothing” was significantly higher rated than “Show automatically” (means 4,35 (STD = 0,86) and 1,82 (STD = 1,01), $t(18) = 7,83$, $p < 0,001$). User preferences, in both web-based and live experiments, were given to the reaction “Ask via mobile device”. The reaction “Ask via mobile” got significantly higher preferences than “Show automatically” (means 4,38 (STD = 0,8) and 2,81 (STD = 1,7), $t(18) = 3,23$, $p = 0,003$). The reaction “Ask via mobile” also got significantly higher preferences than the reaction “Do nothing” (means 4,38 (STD = 0,8) and 2,56 (STD = 1,3), $t(18) = 4,83$, $p < 0,001$). The results reflect the situation with user alone. Similar results were obtained for the situation with spectators present: the reaction “Ask via mobile” got significantly higher user preferences.

In the case of Friend Finder, Scenario 2, the system reaction “Ask via mobile device” got significantly higher rankings than the reaction “Remove automatically” (means 4,35 (STD = 0,86) and 3,59 (STD = 0,87), $t(18) = 2,57$, $p = 0,014$). The reaction “Ask via mobile” also was rated significantly higher than the reaction “Do nothing” (means 4,35 (STD = 0,86) and 1,58 (STD = 0,5), $t(18) = 11,4$, $p < 0,001$). The result refers to the scenario with private data, user alone. A similar result was obtained for the scenario with private data, spectators present: the reaction “Ask via mobile” was rated higher than the reaction “Remove automatically”, though not significantly higher (means 4,05 (STD = 0,9) and 3,56 (STD = 0,9), $t(18) = 1,74$, $p = 0,09$). The reaction “Ask via mobile” also was rated significantly higher than the reaction “Do nothing” (means 4,05 (STD = 0,9) and 1,52 (STD = 0,51), $t(18) = 10,065$, $p < 0,001$). User preferences in the Scenario 2 were given to the reaction “Remove automatically”, in both situations with user alone and present spectators. For example, for the situation with private data, user alone, the reaction “Remove automatically” got significantly higher preferences than the reaction “Ask via mobile device” (means 4,7 (STD = 0,48), $t(18) = 3,23$, $p = 0,003$). Similar results were obtained for the situation with present spectators: the reaction “Do nothing” got significantly higher preferences.

In the case of Travel Planner, Scenario 5, the system reaction “Show on mobile device” got significantly higher rankings in three out of four scenarios: with neutral data, spectators are present, with private data in both cases (user alone or spectators present). For example, in the scenario with private data, user alone, the reaction “Show on mobile device” got significantly higher ratings than “Show on public display” (means 4,31 (STD = 0,6) and 2,56 (STD = 0,51), $t(16) = 8,85$, $p < 0,001$). Similarly, in the scenario with private data, spectator present, the reaction “Show on mobile device” got significantly higher ratings than the reaction “Show on public display” (means 4,19 (STD = 0,54) and 2,31 (STD = 0,6), $t(16) = 9,24$, $p < 0,001$). In the scenario with neutral data, user alone, the participants gave higher rankings to the reaction “Show on public display” than “Show on mobile device” (means 4,38 (STD = 0,5) and 3,125 (STD = 0,61), $t(16) = 6,28$, $p < 0,001$). The high ranking of the reaction “Show on public display” in the latter case was mostly motivated by the convenience of the display. User preferences matched the rankings of the highest trust. In the situation with the neutral data, user alone, the participants gave significantly higher preferences to the reaction “Show on public display”. In the remaining situations the significantly higher preferences were given to the reaction “Show on mobile device”.

Overall, the results indicate that the web-based study provides realistic input for the initialization of the BN despite a few discrepancies.

7.4 Initialization and Validation of the Bayesian Network

In order to evaluate to what extent the system is able to predict the user’s ratings of trust based on her ratings of relevant trust factors (transparency, controllability, comfort of use, seriousness, credibility and security), a model of BN was created using the Genie built-in algorithm for learning Bayesian Networks (see http://genie.sis.pitt.edu/). [5]

Overall, 4 different networks were constructed from the data received in the online studies, one for each scenario shown in the Table 7.2. The first two scenarios were combined into one network. While the basic structure was shared by all networks and was similar to the example network shown in Fig. 7.3, each had different context nodes and display reactions.

Since the evaluations contained quantitative data (rankings of the users), they enabled us to derive distributions of each trust dimension for each contextual combination. For each trust dimension, the author modeled the probability distribution for all combinations of context and display reaction in the BN after the data taken from the online study. The probability distributions for other node combinations, that were not part of the data inquired in the studies (e.g. how Confidence and Competence influence Trust Disposition) were modeled after the results reported in the study [17].

Although the users’ were asked for the preferred display reaction for each context combination as well as their trust in the display reaction presented in each situation, it should be noted that this information was not used to model the network. Instead, the data on user trust estimations and user preferences was used to validate the decisions which were automatically generated by the BN. For the validation of each created network, decisions were generated for all contextual combinations used in the matching scenario of the web-based study. The resulting decisions of the networks were compared to the results from the web-based study. Moreover, the decisions were compared with the user preferences and the estimation of trust.

The contextual combinations were set by entering appropriate evidence into the context nodes, leaving out one node as uncertain. For example, for a specific situation in the Scenario 1, the evidence would be set to “Privacy of Data → Private”, “Movement → Arriving” and “Others Present → uncertain”. Such combination reproduces a real-life situation: often the sensors fail to recognize the presence of the spectators.

[5] The modeling and analysis of the BN was accomplished in a close cooperation with Michael Wissner and Stephan Hammer.

For each combination of context, the display reaction with the highest utility rating was chosen as the system's decision. It referred directly to the computed value of User Trust.

First, the BN generated reactions were compared with user preferences. The comparison was performed between the rankings of the adaptive actions. In other words, the successful result would be the match of the first user-ranked action and the first action generated by the BN.

For each context combination, we selected the display reaction that received the highest average score in the web-based study. When comparing the display reactions preferred by the users with those generated by the respective network, they matched in 19 out of 22 cases (86.36%). In the three mismatching cases, the generated decisions matched the users' second preference.

Second, the BN generated reactions were compared with the users' estimations of their trust. The comparison was performed in the same way as the preferences comparison: the rankings given by the users and the rankings generated by BN were compared. The user rankings were taken as an average of all rankings for the current scenario. The comparison yielded the match in 16 out of 22 cases (72.73%). In two of the mismatching cases, the BN met the second ranking of user trust. In one mismatching case – the third ranking of user trust. Finally, in three cases, the participants' trust ranked two reactions at the exact same position, while the network assigned them slightly differently values, thus preferring one over the other.

These results show that the BN delivers good accuracy in the generated decisions. As an example of this validation, let us take a look at the BN for Scenarios 1 and 2 (see Table 7.2) and its context combinations. Table 7.3 illustrates the matches between the decisions generated by the BN, user preferences, and user estimations of the highest trust. Note that "Adapt" stands for automatic adaptation, "Show data" for "Arriving" context, and "Remove data" for "Leaving" context.

Context			BN Reaction	Preference	Highest Trust
Private	Arriving	Alone	do nothing	ask via mobile	do nothing
Private	Arriving	Spectators	ask via mobile	ask via mobile	do nothing
Private	Leaving	Alone	adapt	adapt	ask via mobile
Private	Leaving	Spectators	adapt	adapt	ask via mobile

Table 7.3. Comparison of the BN generated decisions, user preferences, and user estimations of the highest trust. Example of Friend Finder, Scenario 1 and 2.

It is important to emphasize that the highest user trust does not always match the mostly preferred decision. The results of our web-based and live study support this fact: distributions of user preferences did not always reflect distributions of trust. From the comments of the live study participants, the author found that the feeling of trust often depends on the person's ability to explain the system reaction and agree with it. For example, when a person comes closer to the display, it seems logical and expected that the display does not show any reaction. We learn this behavior from the everyday life: the objects in the interior usually do not react. Obviously, the option "Do nothing" therefore received highest trust rankings. However, the most understandable reaction might not be the most preferred or the most convenient one. Here, the more intelligent (but less expected) reactions were favored. For example, the users found it smart and convenient that the display noticed them and proposed via a mobile device to show their data on the large screen. Thus, the "Ask via mobile device" option was chosen as a preference.

7.5 Summary

This chapter described the mechanism for automatic decision making for adaptive public displays. The mechanism based on probabilistic model, Bayesian Network, enables to estimate user trust for different contextual situations. In a public display environment with a changing social context, user trust is a valuable resource which is critical for user acceptance and satisfaction of the system.

The constructed Bayesian Network was initialized with empirical data collected in web-based study and controlled life experiments. The chapter illustrated the process of the network construction, data collection, and validation of the network.

The presented work aims to inform designers of adaptive displays in the approach to generate adaptation decisions automatically. Trained with empirically collected data, the network can automatically generate decisions on adaptation, best fitting to the current social context. Moreover, the network is able to handle situations when surrounding context is known only partially. The probabilistic nature of the network enables estimations of the best adaptive actions based on the collected knowledge.

Chapter 8

Conclusion and Future Research

8.1 Thesis Outline

The presented work proposes the methods for the context-based adaptation on interactive public displays. In particular, the author focused on social context as a trigger for adaptation. The author motivated the interest to social context as a driver for adaptation by its importance in human-human interaction. Indeed, a changing social context makes people adapt their interaction behaviour, interaction modality, and often the conversation content.

In a similar manner, public displays can be enriched by the ability to sense social context. Such an advance brings the human-display interaction closer to human-human interaction. The context-sensitive displays are able to "understand" their spectators: who is standing in front of the screen, who comes and who leaves, what are the interests or attention level of these people.

Taking a look at the research conducted in the areas of adaptive public displays and context-aware systems, the author identified a lack of profound research on the displays adaptive to social context. The conducted work bridges this gap; it proposes the methods for comprehensive recognition of social context and for automatic display adaptation based on the recognized context. Moreover, the work shows how the displays can take autonomous decisions on adaptation strategy based on the gained knowledge.

The existing works on adaptive displays mostly require spectators to explicitly "announce" their presence in front of the screen. For instance, in order to trigger the adaptation of a screen, the spectators need to activate a Bluetooth client on their mobile devices [6] or send a request to the display [139]. In a real-life public scenario such an explicit adaptation method brings definite limitations. A public display can have spectators who do not possess the required mobile client. Thus, the adaptation will take into account only a part of the spectators equipped with the necessary mobile devices, ignoring the rest of the spectators. In a public place diverse persons are circulating every day. The spectators' constellations therefore can also be diverse and unpredictable: there can be single spectators, groups of spectators, couples, and families. The adaptation must be adjusted not only to the needs of the distinct individuals, but consider the entire group of the spectators.

The current work presents the approach for fully automatic adaptation. The approach is supported by the automatic recognition of social context. The recognition mechanism does not require any explicit input from the spectator side.

Two frameworks were created for the recognition of social context: Ambient Sensing Framework and Mobile Sensing Framework. Both frameworks rely on the autonomous data collection by the means of the sensors. The Ambient Sensing Framework collects the social context with the help of ambient sensors installed in a public place. The framework detects frontal faces of the spectators, their proximity, and conversational activity; it also estimates emotions and the age of the spectators. The framework is able to collect the context data in a multi-display environment. For example, it can trace the social context in a room containing a public display, a tabletop, and mobile devices.

The show cases presented in the thesis demonstrated the applicability of the Ambient Sensing Framework. It can be used to detect the spectators' visual interest and thus adjust the screen content. It can be used to adjust the interaction style with the public display, depending on the number and personalities of people present in the room, as well as on available mobile displays. Finally, it can be used to protect private data of the users, in case an undesired spectator approaches the display.

The Mobile Sensing Framework collects the social context by means of mobile sensors. The sensors integrated into a mobile device collect the data about the device owner. Thus, the framework can detect which locations the user has been visiting, how actively the user moves, how high his or her pulse level is. This data collected over a number of users represents a rich source of social context. It can give detailed information about a social situation within a group of people, a city or even a country. By means of a show case the author demonstrates how the framework can be used to display the social situation within a city going out area. The collected social context is mapped to the social atmosphere within city locations, such as bars, restaurants, cafes. The mobile sensor data collected from the city citizens is transferred to an urban public display, showing the density people in the locations, the loudness, people's motion, and alcohol consumption.

Both frameworks provide automatic collection of the social context, without requiring the user to make any effort for data contribution. Moreover, they allow a flexible extension of the set of the utilized sensors. The set of the sensors presented in the work already delivers a comprehensive picture of the social context. However, the rapidly developing sensor technology is likely to bring new powerful sensors within the next decades. The new sensors can be accommodated into the frameworks, extending the spectrum of the recognized social context.

The knowledge gained in the thesis enabled the author to propose an approach for automatic decision making on the adaptation activity. Following the approach, the displays can decide autonomously how to react on the changing social context. The approach is based on Bayesian network. The structure of the network was designed using experimental data. The network was then initiated with data achieved empirically, and validated against again empirically acquired data. The author shows how the network can be used to generate adaptation decisions, based on the available context data.

8.2 Research Contributions

Below the author summarize the contributions of the thesis. The contributions comprise of technical innovations, theoretical approaches, and software applications. These artefacts can be reused by designers of interactive public displays, as a conceptual inspiration or as a technical guideline.

Ambient Sensing Framework

The framework represents a tool for collection and recognition of social context. It is capable to sense and recognize social context without any contribution from the spectators. The framework delivers a wide range of context data which can be used according to the needs of the public display application. The examples provided in the thesis illustrated how the framework can be used to trace user attention, motion in front of the display, as well as constellations of spectator groups. (Section 3.1)

Mobile Sensing Framework

The Mobile Sensing Framework provides an algorithm for collection of social context by means of mobile sensors. The framework processes the data delivered by the sensors of a modern smartphone: accelerometer, wireless antenna, microphone, gps antenna, gyroscope, and, as a future vision, heart rate sensor. The example provided in the thesis illustrates how the framework can be utilized to collect the social context data from the users in a large scale. (Section 3.2)

The system for automatic adaptation of display content based on group context

By means of the system, the different display content can be mapped to the distinct spectator groups, relying on the group's observation patterns. The content topics, thus, can be tagged as most popular within the certain groups. The tagging is used for further automatic real-time adaptation. Once a certain group is recognized by the display sensors, the content changes for the topic tagged as the most popular for the given group. (Section 4)

Design recommendations for adaptation of interaction technique on public displays

The experimentally derived recommendations consider mobile, physical, and bodily interaction. They explore single- and multi-user scenarios, and describe adaptation guidelines based on the current social context, device context (presence of other displays and mobile devices) and privacy level of the screen content. The recommendations concern the discretion of interaction, physical effort, position freedom, interaction metaphors. They also indicate the critical interaction aspects that have to be avoided, such as unintended logging-in, support of automatic log-out and quick exits. (Section 5.1.7) Design recommendations also explore privacy concerns, space conflicts, and collaboration potential. (Section 5.2.6)

The set of user-defined gestures for interaction on tablet device in a public scenario

The set describes natural gestures which users perform for interaction with tablet in a multi-display environment. The gestures were derived experimentally, exploring three multi-display scenarios: tablet-to-tablet, tablet-tabletop, and tablet-public display. The experiment explains how public setting influences the user choice of gestures. Besides the set of user-defined gestures, the author provides recommendations for the design of iPad gestures, pointing particularly at the physical shape of the devices, inspiration by flat metaphors, and consideration of privacy matters. (Section 5.3.5)

Design recommendations on adaptation strategies for privacy-critical content

The experimentally derived recommendations consider privacy-critical data exposed on a public screen. They explore the combinations of data context (privacy level of the screen content), device context (presence of a mobile display in the setting), and the relationships context (interpersonal relationship between the user and the spectator). For either combination of these context data, the author provides recommendations on screen adaptation strategy. The recommendations map the social relationship of the observer (friend, acquaintance, and stranger) to the desired adaptation technique. The adaptation techniques are recommended depending on the presence of a mobile display in the setting. (Section 6.5)

Bayesian network for automatic decision making on adaptation action

The empirically constructed and initialized network enables to compute the adaptation decision best fitting to the current context. As the input context, the network utilizes social context, device context, and data context. The author illustrates how to construct, initialize, and validate the network. Although the author follows the process for the creation of automatic adaptation on public displays, the similar approach can be followed for the creation of other adaptive systems. (Section 7)

8.3 Future Research Directions

In the thesis, the author takes a profound look into the question of adaptivity on public displays, focusing on adaptive behaviour driven by the changing social context. The author investigated diverse aspects related to the adaptive behaviour: how to sense social context, how to adapt the presentation on the screen and tailor interaction with the screen, how to deal with privacy, how to automate the adaptation.

The next question arising from the conducted work is how to apply the intelligent adaptivity of the displays.

One line of the future research goes in direction of recommender systems. The public displays sensing their spectators can learn the preferences of their audience: recognize their tastes, their wishes, their behavioural patterns. From the learnt knowledge the displays can derive possible recommendations: which content, apart from the learn content, would also interest the spectators. This can be achieved by a recommendation algorithm which analyzes the spectators' preferences. The algorithm can inform the public screen in the content which can be recommended to the current spectator. Thus, the adaptive display, for example, may enhance the sales of a product; and the spectators can profit from more information they would not find otherwise.

Another research line which can benefit from the intelligent adaptivity goes in the direction of persuasive systems. The displays sensitive to social context have potential to influence their spectators. Such displays can directly address the spectator groups or serve as ambient displays conveying a persuasive message to all passing individuals. The example of the first approach is a display calling for better altitude towards elderly. Once the sensors detect a group of youngsters, the screen switches to a campaign calling for elderly support. In the second example, the persuasive message can be addressed to a wide range of spectators. The "Walk & Talk" display presented in the Chapter 3 illustrates such an idea. The display can, for instance, reflect the physical activity of a community of people or their social activity. The display can also aim to change people's behaviour for health reasons: increase awareness of the number of cigarettes smoked today or the number of stairs taken instead of an elevator.

Apart from the application areas of adaptive displays, the future research can explore the scenarios where the users especially profit from automatic adaptation.

One of such scenarios is the in-car scenario. Indeed, t is often critical to know who and in which state is located inside the car. A modern can is equipped with multiple sensors and has diverse displays. The displays include the flat digital screens similar to public displays, as well as panels, glass projections, and light lines. All these types of the displays serve the same purposes: inform the driver about the vehicle state and ease the driving process as much as possible. Design of the displays in the car is indeed a great challenge. On the one hand, the displays must show all necessary information, and in the right time. On the other hand, the displays must stay decent; they should not distract the driver from the main task: driving.

The social situation inside the car vary from very relaxed to very distracting. For example, the driver can be alone, attentive, driving with the day light or at night, or having four passengers who talk to him, or have children who cry and distract him.

A car which is able to recognize the social context could evaluate the situation inside and flexibly adapt the displays to the current social situation. When the driver is alone on a night street, the displays could go down highlighting only the essential information. When the car is full of people who potentially distract the driver, the displays can more prominently present the important data making sure that the driver does not miss it.

Another promising scenario for exploration of adaptive public displays is an urban scene. Multiple examples provided through the thesis show that the urban scene is an exciting playground for deployment of public displays. Urban scenario indeed represents the situation where social context changes frequently and sporadically. The social context which could be extracted from this crowd is diverse: designers could experiment with people's movements, gestures, proximity, and gender. This can be done in absolutely anonymous basis, without intruding individual private sphere, and without extracting any identification of the people.

Although the playground for adaptive displays is quite broad and the technical possibilities and sensors are there, there are still not many adaptive displays in the urban landscape. However, the few existing examples show that such displays can entertain and educate the public, attract interest to the subjects as well as increase the popularity of the buildings "wearing" the displays.

To summarize, adaptive public displays sensitive to social context have potential to enrich not only indoor locations, but can also play important part in in-car scenarios, urban environments, and other areas. This thesis creates the ground for further research on public displays adaptive to social context. The knowledge and materials accumulated by the thesis aim to inform and inspire the researchers and designers in further investigations and development of adaptive public displays.

Author's Publications included into the Thesis

Kurdyukova, E:, Wissner, M., Hammer, S., André, E.: Trust-based Decision-making for the Adaptation of Public Displays in Changing Social Contexts, to appear in *Proceedings of International Conference on Privacy, Security, and Trust*, PST'13, in press (2013).

Kurdyukova, E., Obaid, M., André, E.: Direct, Bodily or Mobile Interaction?: Comparing Interaction Techniques for Personalized Public Displays, *In Proceedings of International Conference on Mobile and Ubiquitous Multimedia*, MOM'12, ACM Press (2012).

Kurdyukova, E., Hammer, S., André, E. Personalization of Content on Public Displays driven by the Recognition of Group Context. In *Proceedings of European Conference on Ambient Intelligence*, AmI'12. Springer Verlag (2012).

Kurdyukova, E., Obaid, M., André, E.: City Pulse: Supporting Going-Out Activities with a Context-Aware Urban Display, In *Proceedings of International Conference on Advances in Computer Entertainment*, ACE'12, Springer Verlag (2012).

Kurdyukova, E., Redlin, M., André, E.: Studying user-defined iPad gestures for interaction in multi-display environment, In *Proceedings of International Conference on Intelligent User Interfaces*, IUI'12, ACM Press (2012).

Kurdyukova, E., André, E., Leichtenstern, K. Trust Management of Ubiquitous Multi-Display Environments. In *Ubiquitous Display Environments*. Springer Verlag (2012).

Kurdyukova, E., Bee, K., André, E. Friend or Foe? Relationship-Based Adaptation on Public Displays. In *Proceedings of European Conference on Ambient Intelligence*, AmI'11. Springer Verlag (2011).

Kurdyukova, E. André, E.: Designing Transparent and Controllable Adaptation to Social Context. In *Adjunct Proceedings of international conference on User Modeling, Adaptation and Personalization*, UMAP'11, Springer Verlag (2011).

Kurdyukova, E.: Designing Trustworthy Adaptation on Public Displays, In *Proceedings of International Conference on User Modeling, Adaptation and Personalization*, UMAP'11, Springer Verlag (2011).

Kurdyukova, E., André, E., Leichtenstern, K.: A Decision-Theoretic Approach to Maintain Trust in Ubiquitous Display Environments, In *International Journal of Computing*, Volume 9 (3), 2010.

Kurdyukova, E. André, E., Leichtenstern, K.: Trust-centered Design for Multi-Display Applications. In *Proceedings of International Conference on Advances in Mobile Computing & Multimedia*, MoMM'10, ACM Press (2010).

References

1. Abowd, G. D., Mynatt, E. D.: Charting past, present and future research in ubiquitous computing. In *ACM Transactions on Computer-Human Interactions*, 7(1), 2001.
2. Ackerman, M.S., Mainwaring, S.D.: *Privacy Issues and Human-Computer Interaction. Security and Usability: Designing Secure Systems That People Can Use*. O'Reilly (2005).
3. Agre, P. Changing places: Contexts of awareness in computing. In *Human Computer Interaction*. 16, (2001).
4. Aiello, J.R.: Human spatial behavior. In *Handbook of environmental psychology*. New York: John Wiley & Sons (1987).
5. Alexander, C., Ishikawa, S., Silverstein, M., Jacobson, M., Fiksdahl-King, I., Angel, S.: *A Pattern Language: Towns, Buildings, Construction*. Oxford University Press (1977).
6. Alt, F., Balz, M., Kristes, S., Shirazi, A.S., Mennenöh, J., Schmidt, A., Schröder, H., Goedicke, M.: Adaptive User Profiles in Pervasive Advertising Environments. In *Proceedings of European Conference on Ambient Intelligence*, AmI'09. Springer Verlag (2009).
7. AppleInsider: www.appleinsider.com/articles/09/11/19/apple_investigates_space_age_fitness_tracking_technology.html
8. Apple iPod: http://www.apple.com/en/ipod/
9. Ballagas, R., Rohs, M., Sheridan, J. G.: Sweep and point and shoot: phonecam-based interactions for large public displays. In *Proceedings of International Conference on Human Factors in Computer Systems*, CHI'05, ACM Press (2005).
10. Baldauf, M., Fröhlich, P. The Augmented Video Wall: Multi-user AR Interaction With Public Display. In *Proceedings of International Conference on Human Factors in Computer Systems*, CHI'13, ACM Press (2013).
11. Barkhuus, L., Dey, A.K.: Is Context-awareness computing taking control away from the user? Three levels of interactivity examined. In *Proceedings of International Conference on Ubiquitous Computing*, UbiComp'03. Springer-Verlag (2003).
12. Bardram, J. E., Norskov, N.: A context-aware patient safety system for the operating room. In *Proceedings of International Conference on Ubiquitous Computing*, UbiComp'04. ACM Press (2004).
13. Bardram, J., Friday, A.: Ubiquitous Computing Systems. In Ubiquitous Computing Fundamentals. CRC Press (2009).

14. Baudisch, P., DeCarlo, D., Duchowski, A., Geisler, B.: Focusing on the Essential: Considering Attention in Display Design. In *Communications of the ACM,* 46(3), March 2003.

15. Ballagas, R., Rohs, M., Sheridan, J., Borchers, J.: The Design Space of Mobile Phone Input Techniques for Ubiquitous Computing. In *Handbook of Research on User Interface Design and Evaluation for Mobile Technologies*, IGI Global (2008).

16. Baudisch, P., Cutrell, E., Robbins, D., Czerwinski, M., Tandler, P., Bederson, B., Zierlinger, A.: Drag-and-pop and drag-and-pick: Techniques for accessing remote screen content on touch- and pen-operated systems. In *Proceedings of International Conference on Human-Computer Interaction*, INTERACT'03, Springer Verlag (2003).

17. Bee, K., Hammer, S., Pratsch, C., Andre, E.: The automatic trust management of self-adaptive multi-display environments. In *Trustworthy Ubiquitous Computing*, Atlantis Press, Volume 6 (2012).

18. Bellotti, V., Edwards, W.K.: Intelligibility and accountability: human considerations in context-aware systems. In *Human-Computer Interaction*, 16, 2 (2001).

19. Bergstrom, T., Karahalios, K.: Conversation Votes: Enabling Anonymous Cues. In *Extended Abstracts on Human Factors in Computing Systems*, CHI EA'07, ACM Press (2007).

20. Bezerianos, A., Balakrishnan, R.: The Vacuum: Facilitating the Manipulation of Distant Objects. In *Proceedings of International Conference on Human Factors in Computing Systems*, CHI 2005, ACM Press (2005).

21. Bianchi-Berthouze, N., Kim, W. W., Patel, D. Does Body Movement Engage You More in Digital Game Play? and Why? In *Proceedings of International Conference on Affective Computing and Intelligent Interaction*, ACII'07, Springer Verlag (2007).

22. Blinkenlights Project, blinkenlights.net.

23. Borchers, J., Lee, E., Samminger, W., Mühlhäuser, M.: Personal Orchestra: A real-time audio/video system for interactive conducting. In *ACM Multimedia Systems*, 9(5), March 2004.

24. Boring, S., Baur, D., Butz, A., Gustafson, S., Baudisch, P. Touch projector: mobile interaction through video. In *Proceedings of International Conference on Human Factors in Computing Systems*, CHI'10, ACM Press (2010).

25. Boring, S., Gehring, S., Wiethoff, A., Blöckner, M., Butz, A.: Multi-User Interaction on Media Facades through Live Video on Mobile Devices. In *Proceedings of International Conference on Human Factors in Computing Systems*, CHI'11. ACM Press (2011).

26. Boring, S., Jurmu, M., and Butz, A.: Scroll, tilt or move it: using mobile phones to continuously control pointers on large public displays, In *Proceedings of Australasian Conference on Human Factors in Computing Systems*, ACM Press (2009).

27. Boyle, M., Edwards, C., Greenberg, S.: The Effects of Filtered Video on Awareness and Privacy, In: *Proceedings of the International Conference on Computer supported cooperative work*, CSCW'00, ACM Press (2000).

28. Bradley, N. A., Dunlop, M. D.: Understanding contextual interactions to design navigational context-aware applications. In *Proceedings of International Conference on Human-Computer Interaction with Mobile Devices and Services*, Mobile HCI'02. Springer Verlag (2002).

29. Brignull, H., Izadi , S., Fitzpatrick, G., Rogers, Y., Rodden, T.: The introduction of a shared interactive surface into a communal space, In *Proceeding of International Conference on Computer supported cooperative work*, CSCW'04, ACM Press (2004).

30. Broll, G., Hausen, D.: Mobile and Physical User Interfaces for NFC-based Mobile Interaction with Multiple Tags, In *Proceedings of International Conference on Human-Computer Interaction with Mobile Devices and Services*, MobileHCI'10, ACM Press (2010).

31. Broll, G., Reithmeier, W., Holleis, P., and Wagner, M. Design and Evaluation of Techniques for Mobile Interaction with Dynamic NFC-Displays. In *Proceedings of International Conference on Tangible and Embedded Interaction*, TEI'11, ACM Press (2011).

32. Broll, G., Siorpaes, S., Rukzio, E., Paolucci, M., Hamard, J., Wagner, M., Schmidt, A.: Comparing Techniques for Mobile Interaction with Objects from the Real World, *International Workshop on Pervasive Mobile Interaction Devices* (2007).

33. Buxton, W.: Integrating the Periphery and Context: A New Model of Telematics. In *Proceedings of Graphics Interface*, (1995).

34. Cao, H., Olivier, P., Jackson, D.: Enhancing privacy in public spaces through crossmodal displays. In *Social Science Computer Review*, 26(1), (2008).

35. Chapman, A. J.: Eye Contact, Physical Proximity, and Laughter: A Re-examination of the Equilibrium Model of Social Intimacy. In *Social Behavior and Personality*, Volume 3 (2), 1975.

36. Cheverst, K., Byun, H.E., Fitton, D., Sas. C., Kray, C., Villar, N.: Exploring Issues of User Model Transparency and Proactive Behaviour in an Office Environment Control System, In *User Modeling and User-Adapted Interaction*, 15 (3-4), 2005.

37. Cheverst, K., Davies, N., Mitchell, K., Efstratiou, C.: Using context as a crystal ball: rewards and pitfalls. In Personal Ubiquitous Computing, 5 (1), 2001.

38. Cheverst, K., Dix, A., Fitton, D., Kray, C., Rouncefield, M., Sas, C., Saslis-Lagoudakis, G., Sheridan, J.: Exploring Bluetooth based Mobile Phone Interaction with the Hermes Photo Display. In *Proceedings of International Conference on Human-Computer Interaction with Mobile Devices and Services*, MobileHCI'05, ACM Press (2005).

39. Churchill, E., Girgensohn, A., Nelson, L., Lee, A.: Blending digital and physical spaces for Ubiquitous community participation. In *Communications of the ACM*, Volume 47 (2), ACM Press (2004).

40. Churchill, E.F., Nelson, L., Hsieh, G.: Cafe Life in the Digital Age: Augmenting Information Flow in a cafe-work-entertainment space. In *Extended Abstracts on Human Factors in Computing Systems*, CHI EA'06, ACM Press (2006).

41. Churchill, E., Nelson, L., Denoue, L., Girgensohn, A.: The Plasma Poster Network: Posting Multimedia Content in Public Places, In *Proceedings of International Conference on Human-Computer Interaction*, INTERACT'03, Springer Verlag (2003).

42. Churchill, E., Nelson, L., Denoue, L., Murphy, P., Helfman, J.: The Plasma Poster Network. In Public and Situated Displays. Kluver Academic Publishers (2003).

43. Cockburn, A., Karlson, A., Bederson, B.: A Review of Overview+Detail, Zooming, and Focus+Context Interfaces, In *ACM Computing Surveys*, Volume 41 (1), ACM Press (2008).

44. Congleton, B., Ackerman, M., Newman, M.: The ProD Framework for Proactive Displays, In *Proceedings of ACM Symposium on User Interface Software and Technology*, UIST'08, ACM Press (2008).

45. Crabtree, B., Rhodes, B.: Wearable computing and the remembrance agent. In *BT Techno*, 16 (3), 1998.

46. Dey, A. K.: Understanding and using context. In *Personal Ubiquitous Computing*, 5, (1), 2001.

47. Dey, A. K., Newberger, A.: Support for context-aware intelligibility and control. In *Proceedings of International Conference on Human Factors in Computing Systems*, CHI'09, ACM Press (2009).

48. De Vries, P., Midden C.: The effects of error son system trust, self-confidence, and the allocation of control in route planning. In *International Journal of Human-Computer Studies*, 58 (2003).

49. Dietz, P., Raskar, R., Booth, S., van Baar, J., Wittenburg, K., Knep, B.: Multi-Projectors and Implicit Interaction in Persuasive Public Displays, In *Proceedings of the working conference on Advanced Visual Interfaces*, AVI'04, ACM Press (2004).

50. Diaz-Marino, R., Greenberg, S.: The proximity toolkit and viconface. In *Extended Abstracts of international conference on Human Factors in Computing Systems*, CHI EA'10, ACM Press (2010).

51. Dourish, P. Seeking a foundation for context-aware computing. In *ACM Transactions on Human-Computer Interaction*, Volume 16, ACM Press (2001)

52. Dourish, P. What we talk about when we talk about context. In *Personal Ubiquitous Computing*, 8 (1), 2004.

53. Elting, C.: Orchestrating output devices. In *Proceedings of the joint conference on Smart objects and ambient intelligence innovative context-aware services: usages and technologies*, ACM Press (2005).

54. Exeler, J., Buzeck, M., Müller, J.: eMir: Digital Signs that react to Audience Emotion. In *Proceedings of the Workshop on Pervasive Advertising*, (2009).

55. Facebook. https://www.facebook.com/

56. Fass, A., Pausch, R.: Adding Scripting to a Public Bulletin Board. In *Proceedings of ACM Symposium on User Interface Software and Technology*, UIST'02, ACM Press (2002).

57. Fisher, D.: Hotmap: Looking at Geographic Attention, In IEEE Transactions on Visualization and Computer Graphics, Volume 13 (6), 2007.

58. Finke, M., Tang, A., Leung, R., Blackstock, M.: Lessons Learned: Game Design for Large Public Displays, In *Proceeding of international conference on Digital Interactive Media in Entertainment and Arts*, DIMEA'08, ACM Press (2008).

59. Fitzmaurice G.: Situated Information Spaces and Spatially Aware palmtop computers, In *Communications of the ACM*, Volume 36 (7), ACM Press (1993).

60. Flickr. http://www.flickr.com/

61. Fraser, K., Rodden, T., O'Malley, C.: Trust, Privacy and Relationships in "Pervasive Education": Families' View on Homework and Technologies, In *Proceeding of International Conference on Pervasive Computing*, Pervasive'07, Springer Verlag (2007).

62. Foursquare. https://de.foursquare.com/

63. Froehlich, J. Neumann, J. Oliver, N.: Sensing and Predicting the Pulse of the City through Shared Bicycling, In *Proceedings of International Joint Conference on Artificial Intelligence*, IJCAI'09, Morgan Kaufmann Publishers Inc. (2009).

64. Fujita, K., Itoh, Y., Ohsaki, H., Ono, H., Kagawa, K., Takashima, K., Tsugawa. S., Nakajima, K., Hayashi, Y., Kishino, F.: Ambient Suite: enhancing communication among multiple participants, In *Proceeding of International Conference on Advances in Computer Entertainment Technology*, ACE'11, ACM Press (2011).

65. Gilroy, S. W., Cavazza, M., Chaignon, R., Mäkeläa, S., Niiranen, M., André, E., Vogt, T., Billinghurst, M., Seichter, H., Benayoun, M.: An emotionally responsive AR art installation. In *Proceedings of the Workshop on Mixed Reality Entertainment and Art*, (2007).

66. Girardin, F., Calabrese, F., Dal Fiore, F., Ratti, C., Blat, J.: Digital Footprinting: Uncovering Tourists with User-Generated Content. In *IEEE Pervasive Computing*, Volume 7 (4), October 2008.

67. Girardin, F., Fiore, F. D., Blat, J., Ratti, C.: Understanding of tourist dynamics from explicitly disclosed location information. In *Proceedings of the International Symposium on Location-based Services and Telecartography* (2007).

68. Girardin, F., Fiore, F. D., Ratti, C., Blat, J.: Leveraging explicitly disclosed location information to understand tourist dynamics: a case study. In *Journal of Location-Based Services*, Volume 2 (1), March 2008.

69. Glass, A., McGuinness, D.L., Wolverton. M.: Toward establishing trust in adaptive agents. In *Proceedings of International Conference on Intelligent User Interfaces*, IUI'08, ACM Press (2008).

70. Graham C., Cheverst. K.: Guides, locals, chaperones, buddies and captains: managing trust through interaction paradigms. In *Proceedings of Symposium on HCI with Mobile Devices and Services*, ACM Press (2004).

71. Graham, C., Cheverst, K., Rouncefield, M., Kray, C.: Going more public: Situated Display design in a care setting through co-realization, In *Proceedings of conference on Designing for User eXperience*, DUX '05, AIGA: American Institute of Graphic Arts (2005).

72. Grandison, T., Sloman, M.: A Survey of Trust in Internet Applications. In *Communications Surveys and Tutorials*, Volume 3 (4), 2000.

73. Greaves, A., Rukzio, E.: View & Share: Co-Present Viewing and Sharing of Pictures using Personal Projectors. In *International Journal of Mobile Human Computer Interaction*, (2010).

74. Greenberg, S.: Opportunities for Proxemic Interactions in Ubicomp, Keynote at International Conference on Human Computer Interaction, INTERACT'11, (2011).

75. Greenberg, S., Rounding, M.: The Notification Collage: Posting Information to Public and Personal Displays, In *Proceedings of the International Conference on Human Factors in Computing Systems*, CHI'01, ACM Press (2001).

76. Greenberg, S., Marquardt, N., Ballendat, T., Diaz-Marino, R., Wang, M.: Proxemic Interactions: The New Ubicomp? In *ACM Interactions*, 18(1), ACM Press (2011).

77. Goodchild, M.F.: Citizens as voluntary Sensors: Spatial Data Infrastructure in the World of Web 2.0. In *International Journal of Spatial Data Infrastructures Research*. Volume 2, 2007.

78. Guye-Vuillème, A.: Simulation of Nonverbal Social Interaction and small groups Dynamics in virtual Environments, Ph.D. Thesis, Lausanne, 2004.

79. Hall, E.T.: Proxemics, In *Current Anthropology*, Volume 9 (2/3), June 1968.

80. Hall, E.T.: *The Hidden Dimension*, Anchor Books (1990).

81. Hardy, R., Rukzio, E., Holleis, P., Wagner, M.: Mobile Interaction with Static and Dynamic NFC-based Displays, In *Proceedings of International Conference on Human-Computer Interaction with Mobile Devices and Services*, MobileHCI'10, ACM Press (2010).

82. Harrison, C., Dey, A.K.: Lean and Zoom; Proximity-Aware User interface and content magnification, In *Proceedings of the International Conference on Human Factors in Computing Systems*, CHI'08, ACM Press (2008).

83. Hawkey, K., Kellar, M., Reilly, D., Whalen, T., Inkpen, K.: The proximity factor: impact of distance on co-located collaboration, In *Proceedings of international ACM conference on Supporting group work*, GROUP'05, ACM Press (2005).

84. HearChamberOrchestra Project: http://heartchamberorchestra.org

85. Holleis, P., Rukzio, E., Otto, F., Schmidt, A.: Privacy and Curiosity in Mobile Interactions with Public Displays, In *Adjunct Proceedings of International Conference on Human Factors in Computing Systems*, CHI'07, ACM Press (2007).

86. Horvitz, E., Jacobs, A., Hovel, D.: Attention-Sensitive Alerting. In *Proceedings of Conference on Uncertainty in Artificial Intelligence*, UAI'99, Morgan Kaufmann (1999).

87. Horvitz, E., Kadie, C., Paek, T., Hovel, D.: Models of Attention in Computing and Communication: From Principles to Applications. In *Communications of the ACM*, Volume 46 (3), 2003.

88. Horvitz, E., Koch, P., Kadie, C., Jacobs, A.: Coordinate: Probabilistic Forecasting of Presence and Availability. In *Proceedings of Conference on Uncertainty in Artificial Intelligence*, UAI'02, Morgan Kaufmann (2002).

89. Houghton, D.J., Joinson, A.N.: Privacy, Social Network Sites, and Social Relations, In *Journal of Technology in Human Services*, Volume 28 (1-2), 2010.

90. Huang, E., Koster, A., Borchers, J.: Overcoming assumptions and uncovering practices: When does the public really look at public displays? In *Proceedings of International Conference on Pervasive Computing*, Pervasive'08. Springer Verlag (2008).

91. Huang, E., Mynatt, E.: Semi-Public Displays for Small, Co-located Groups, In *Proceedings of international conference on Human factors in computing systems*, CHI'03, ACM Press (2003).

92. Iachello, G., Smith, I., Consolvo, S., Chen, M., Abowd, G.: Developing Privacy Guidelines for Social Location Disclosure Applications and Services, In *Proceedings of the symposium on Usable privacy and security*, SOUPS '05, ACM Press (2005).

93. IMEC Research, www.imec.be/be_en/research.html

94. Interactive Video Wall of Copenhagen, http://museummedia.nl/2011/04/museum-of-copenhagen-interactive-video-wall-housed-in-a-shipping-container/

95. Jacucci, G., Spagnolli, A., Chalambalakis, A., Morrison, A., Liikkanen, L., Roveda, S., Bertoncini, M.: Bodily Explorations in Space: Social Experience of a Multimodal Art Installation, In *Proceedings of the International Conference on Human-Computer Interaction*, INTERACT'09, Springer Verlag (2009).

96. Jameson, A., Smyth, B.: Recommendation to groups. In T*he Adaptive Web: Methods and Strategies of Web Personalization*, Springer Verlag (2007).

97. Jannach, D., Zanker, M., Felfernig, A., Friedrich, G.: *Recommender Systems: An Introduction*. Cambridge University Press (2010).

98. John, L., Rist, T.: xioScreen: Experiences Gained from Building a Series of Prototypes of Interactive Public Displays. In *Ubiquitous Display Environments*. Springer Verlag (2012).

99. Jones, S., O'Neill, E.: Feasibility of Structural network clustering for group-based privacy control in social networks, In *Proceedings of Symposium on Usable Privacy and Security*, SOUPS'10, ACM Press (2010).

100. Ju, W. G., Lee, B. A., Klemmer, S. R.: Range: exploring proxemics in collaborative whiteboard interaction. In *Extended Abstracts on Human Factors in Computing Systems*, CHI EA'07, ACM Press (2007).

101. Karam, M., Payne, T., David, E.: Evaluating BluScreen: Usability for Intelligent Pervasive Displays. In *Proceedings of International Conference on Pervasive Computing and Applications*, Pervasive'07, IEEE Press (2007).

102. Kjeldskov J., Paay J.: Indexicality: understanding mobile human-computer interaction in context. In *ACM Transactions on Computer-Human Interaction*, 17(4), AVM Press (2010).

103. Microsoft Kinect: http://www.xbox.com/en-US/kinect

104. Kini, A., Choobineh, J.: Trust in electronic commerce: definition and theoretical considerations. In *Proceedings of Hawaii International Conference on System Sciences*, IEEE Computer Society (1998).

105. Kisilevich, S., Mansmann, F., Bak, P., Keim, D., Tchaikin, A.: Where would you go on your next vacation? A framework for visual exploration of Attractive Places. In *Proceedings of the International Conference on Advanced Geographic Information Systems, Applications, and Services*, IEEE Computer Society (2010).

106. Kray, C., Nesbitt, D., Dawson, J., Rohs, M.: User-defined gestures for connecting mobile phones, public displays, and tabletops. In *Proceedings of International Conference on Human-Computer Interaction with Mobile Devices and Services*, MobileHCI'11. ACM Press (2011).

107. Kurdyukova, E. André, E.: Designing Transparent and Controllable Adaptation to Social Context. In *Adjunct Proceedings of international conference on User Modeling, Adaptation and Personalization*, UMAP'11, Springer Verlag (2011).

108. Kurdyukova, E., André, E., Leichtenstern, K.: Introducing Multiple Interaction Devices to Interactive Storytelling: Experiences from Practice, In *Proceedings of International Conference on Interactive Digital Storytelling*, ICIDS'09, Springer Verlag (2009).

109. Kurdyukova, E. André, E., Leichtenstern, K.: Trust-centered Design for Multi-Display Applications. In *Proceedings of international conference on Advances in Mobile Computing & Multimedia*, MoMM'10, ACM Press (2010).

110. Kurdyukova, E., André, E., Leichtenstern, K. Trust Management of Ubiquitous Multi-Display Environments. In *Ubiquitous Display Environments*. Springer Verlag (2012).

111. Kurdyukova, E., Bee, K., André, E. Friend or Foe? Relationship-Based Adaptation on Public Displays. In *Proceedings of European Conference on Ambient Intelligence*, AmI'11. Springer Verlag (2011).

112. Kurdyukova, E., Hammer, S., André, E. Personalization of Content on Public Displays driven by the Recognition of Group Context. In *Proceedings of European Conference on Ambient Intelligence*, AmI'12. Springer Verlag (2012).

113. Kurdyukova, E., Redlin, M., André, E.: Studying user-defined iPad gestures for interaction in multi-display environment, In *Proceedings of International Conference on Intelligent User Interfaces*, IUI'12, ACM Press (2012).

114. Küblbeck, C., Ernst, A.: Face detection and tracking in video sequences using the modified census transformation. In *Image and Vision Computing,* 24 (6), 2006.

115. Kwapisz, J.R., Weiss, G.M., Moore, S.A.: Activity recognition using cell phone accelerometers. In *ACM Explorations Newsletter*, Volume 12 (2), ACM Press (2010).

116. Lalintha Peiris, R., Newton Fernando, O.N., Cheok, A.D.: Flexible, Non-emissive Textile Display. In *Proceedings of European Conference on Ambient Intelligence*, AmI'11. Springer Verlag (2011).

117. Lamming, M. G., Newman, W. M.: Activity-based information retrieval: technology in support of personal memory. In *Proceedings of the World Computer Congress on Personal Computers and Intelligent Systems* (Information Processing). North-Holland Publishing (1992).

118. Langheinrich, M.: A Privacy Awareness System for Ubiquitous Computing Environments, In *Proceedings of International Conference on Ubiquitous Computing*, UbiComp'02. Springer Verlag (2002).

119. Langheinrich, M.: Privacy by Design – Principles of Privacy-Aware Ubiquitous Systems. In *Proceedings of International Conference on Ubiquitous Computing*, UbiComp'01. Springer Verlag (2001).

120. Langheinrich, M. Privacy in Ubiquitous Computing. In *Ubiquitous Computing Fundamentals*. CRC Press (2010).

121. Lee, J., Lee, J., Kim, H., and Kim, J. Gesture-Based Interactions on Multiple Large Displays with a Tabletop Interface, In *Proceedings of International Conference on Universal Access in Human-Computer*, UACHI'07, Springer Verlag (2007).

122. Lee, H., Jeong, H., Lee, J., Yeom, K., Park, J.: Gesture-Based Interface for Connection and Control of Multi-device in a Tabletop Display Environment, In *Proceedings of International Conference on Human-Computer Interaction*, HCI'09, Springer Verlag (2009).

123. Liikkanen, L., Jacucci, G., Huvio, E., Laitinen, T., Andre, E.: Exploring Emotions and Multimodality in Digitally Augmented Puppeteering, In *Proceedings of the working conference on Advanced visual interfaces*, AVI'08. ACM Press (2008).

124. Lim, B.Y., Dey, A.K.: Toolkit to Support Intelligibility in Context-Aware Applications. In *Proceedings of the International Conference on Ubiquitous Computing*, UbiComp'10, Springer Verlag (2010).

125. Lim, B.Y., Dey, A.K.: Why and Why not explanations improve the intelligibility of context-aware intelligent systems. In *Proceedings of the International Conference on Ubiquitous Computing*, UbiComp'10, Springer Verlag (2010).

126. Lin, J.J., Mamykina, L., Lindtner, S., Delajoux, G., Strub, H.: Fish'n'Steps: Encouraging Physical Activity with an Interactive Computer Game. In *Proceedings of International Conference on Ubiquitous Computing*, UbiComp‘06. Springer-Verlag (2006).

127. Lindley, S. E., Le Couteur, J., Berthouze, N. L.: Stirring up experience through movement in game play: effects on engagement and social behaviour. In *Proceedings of International Conference on Human Factors in Computing Systems*, CHI’08, ACM Press (2008).

128. Little, L., Briggs, P.: Private whispers/public eyes: Is receiving highly personal information in a public place stressful? In *Interacting with Computers*, Volume 21 (4), August 2009.

129. Lucero, A., Aliakseyeu, D., Martens, J.: Funky wall: presenting mood boards using gesture, speech and visuals, In *Proceedings of the working conference on Advanced visual interfaces*, AVI '08, ACM Press (2008).

130. Lumsden, J., MacKay, L.: How does personality affect trust in B2B e-commerce?, In *Proceedings of the International conference on Electronic commerce*, ICEC’06, ACM Press (2006).

131. Mahato, H., Kern. D., Holleis. P., Schmidt, A.: Implicit Personalization of Public Environments using Bluetooth. In *Proceedings of International Conference on Human factors in computing systems*, CHI’08, ACM Press (2008)

132. Marquardt, N., Diaz-Marino, R., Boring, S., Greenberg, S.: The Proximity Toolkit: Prototyping Proxemic Interactions in Ubiquitous Computing Ecologies. In *Proceedings of ACM Symposium on User Interface Software and Technology*, UIST'11, ACM Press (2011).

133. Marquardt, N., Greenberg, S.: Informing the Design of Proxemic Interactions. In *Pervasive Computing*, Volume 11 (2), 2012.

134. Masthoff, J.: Group Modeling: Selecting a Sequence of Television Items to Suit a Group of Viewers. In *User Modeling and User-Adapted Interaction,* 14(1), 2004.

135. Mathew, A., Rogers, Y., Lloyd, P.: Post-it note art: evaluating public creativity at a user generated art installation. In *Proceedings of ACM conference on Creativity and cognition*, ACM Press (2011).

136. McCarthy, J.F., Farnham, S., Patel, Y., Ahuja, S., Norman, D., Hazlewood, W., Lind, J.: Supporting Community in Third Places with Situated Social Software. In *Proceedings of International Conference on Communities and Technologies*, C&T’09, ACM Press (2009).

137. McCarthy, J., McDonald, D., Soroczak, S., Nguyen, D., Rashid, A.: Augmenting the Social Space of an Academic Conference. In *Proceedings of the ACM conference on Computer supported cooperative work*, CSCW’04, ACM Press (2004).

138. McCullough, M.: *Digital Ground – Architecture, Pervasive Computing, and Environmental Knowing*. MIT Press (2004).

139. Memarovic, N., Elhart, I., Langheinrich, M.: FunSquare: First Experiences with Autopoiesic Content. In *Proceedings of the International Conference on Mobile and Ubiquitous Multimedia*, MUM'11, ACM Press (2011).

140. Meschtscherjakov, A., Reitberger, W., Mirlacher, T., Huber, H., Tscheligi, M.: AmIQuin – An Ambient Mannequin for the Shopping Environment. In *Proceedings of European Conference on Ambient Intelligence,* AmI'09, Springer Verlag (2009).

141. Mullen B.: Self-attention theory: the effects of group composition on the individual. In *Springer Series in Social Psychology* (1987).

142. Müller, J., Alt, F. Michelis, D.: *Pervasive Advertising*, Springer Verlag (2011).

143. Müller, J., Alt, F., Michelis, D., Schmidt, A.: Requirements and design space for interactive public displays. In *Proceedings of ACM Conference on Multimedia*, ACM Press (2010).

144. Müller, J., Exeler, J., Buzeck, M., Krüger, A.: ReflectiveSigns: Digital Signs that Adapt to Audience Attention. In *Proceedings of International Conference on Pervasive Computing*, Pervasive'09, Springer Verlag (2009).

145. Müller, J., Krüger, A.: MobiDiC: Context Adaptive Digital Signage with Coupons. In *Proceedings of European Conference on Ambient Intelligence*, AmI'09, Springer Verlag (2009).

146. Müller, J., Krüger, A., Kuflik, T.: Maximizing the Utility of Situated Public Displays. In *Adjunct Proceedings of International Conference on User Modeling*, UM'07, Springer Verlag (2007).

147. Müller, J., Paczkowski, O., Krüger, A.: Situated Public News & Reminder Displays. In *Proceedings of European Conference on Ambient Intelligence*, AmI'07, Springer Verlag (2007).

148. Müller, J., Walter, R., Bailly, G., Nischt, M., Alt, F.: Looking Glass: A Field Study on Noticing Interactivity of a Shop Window. In *Proceedings of International Conference on Human Factors in Computing Systems*, CHI'12, ACM Press (2012).

149. Müller, J., Wilmsmann, D., Exeler, J., Buzeck, M., Schmidt, A., Jay, T., Krüger, A.: Display blindness: The effect of expectations on attention towards digital signage. In *Proceedings of International Conference on Pervasive Computing*, Pervasive'09, Springer Verlag (2009).

150. NAudio Open Source .Net Library: http://naudio.codeplex.com/

151. Nielsen, J.: *Usability Engineering*. Morgan Kaufmann (1993).

152. Ning, T., Müller, J., Walter, R., Bailly, G., Wacharamanotham, C., Borchers, J., Alt, F.: No Need To Stop: Menu Techniques for Passing by Public Displays. In *Workshop on Large Displays in Urban Life at International Conference on Human Factors in Computing Systems* (2011).

153. Norman, D.: *The Design of Everyday Things*, Basic Books (2002).

154. Obaid, M., Kurdyukova, E., André, E.: City Pulse: Supporting Going-Out Activities with a Context-Aware Urban Display, In *Proceedings of International Conference on Advances in Computer Entertainment*, ACE'12, Springer Verlag (2012).

155. Ojala, T., Kostakos, V., Kukka, H., Heikkinen, T., Lindén, T., Jurmu, M., Hosio, S.: Multipurpose interactive public displays in the Wild: Three years later. *IEEE Computer* 45(5), 2006.

156. O'Hara, K., Glancy, M., Robertshaw, S.: Understanding collective play in an urban screen game. In *Proceedings of the ACM conference on Computer supported cooperative work*, CSCW'08, ACM Press (2008).

157. O'Hara, K., Perry, M., Churchill, E.: *Public and Situated Displays: Social and Interactional Aspects of Shared Display Technologies*, Kluwer Academic Publishers, (2004).

158. O'Neill, E., Kostakos, V., Kindberg, T., Fatah gen. Schiek, A., Penn, A., Stanton Fraser, D. Jones, T.: Instrumenting the City: Developing Methods for Observing and Understanding the Digital Cityscape. In *Proceedings of the International Conference on Ubiquitous Computing*, UbiComp'06, Springer Verlag (2006).

159. OpenCV software library. Open Source Computer Vision. http://opencv.org/

160. O'Reilly, T. What is Web 2.0? Design Patterns and Business Models for the Next Generation of Software (2005).

www.oreillynet.com/pub/a/oreilly/tim/news/2005/09/30/what-is-web-20.html

161. Osmond, H.: *Function as the basis of psychiatric ward design*. Psychiatric Services (1957).

162. Paay, J., Kjeldskov, J., Howard, S., Dave, B.: Out on the town: A socio-physical approach to the design of a context aware urban guide. In *Transactions on Computer-Human Interaction*, 16 (2), 2009.

163. Palen, L., Dourish, P.: Unpacking "Privacy" for Networked World, In *Proceedings of the International conference on Human factors in computing systems*, CHI'03, ACM Press (2003).

164. Pease, A., Pease, B.: *The Definitive Book of Body Language: How to Read Others' Attitudes by Their Gestures*, Sheldon Press (1988).

165. Peltonen, P., Salovaara, A., Jacucci, G., Ilmonen, T., Ardito, C., Saarikko, P., Batra, V.: Extending large-scale event participation with user-created mobile media on a public display. In *Proceedings of International conference on Mobile and Ubiquitous Multimedia*, MUM'07, ACM Press (2007).

166. Pering, T., Ballagas, R., Want, R.: Spontaneous marriages of mobile devices and interactive spaces. In *Communications of the ACM*, Volume 48 (9), ACM Press (2005).

167. Pulse Room Installation: www.lozano-hemmer.com/pulse_room.php

168. Rashid, A.M., Albert, I., Cosley, D., Lam, S.K., McNee, S.M., Konstan, J.A., Riedl, J.: Getting to know you: learning new user preferences in recommender systems. In *Proceedings of International Conference on Intelligent User Interfaces*, IUI'02, ACM Press (2002).

169. Rashid, U., Nacenta, M., Quigley, A.: The Cost of Display Switching: A Comparison of Mobile, Large Display and Hybrid UI Configurations, In *Proceedings of the working conference on Advanced visual interfaces,* AVI '12, ACM Press (2012).

170. Rantanen, J., Impio, J., Karinsalo, T., Reho, A., Tasanen, M., Vanhala, J.: Smart clothing prototype for the arctic environment. In *Personal Ubiquitous Computing*, 6 (1), 2002.

171. Ratti, C., Pulselli, R.M., Williams, S., Frenchman, D.: Mobile Landscapes: Using Location Data from Cell Phones for Urban Analysis. In *Environment and Planning*, Volume 33 (5), 2006.

172. Ravi, N., Dandekar, N., Mysore, P., Littman, M. L.: Activity recognition from accelerometer data. In *Proceedings of the conference on Innovative applications of artificial intelligence*, IAAI'05, AAAI Press (2005).

173. Reeves, B., Nass, C.: *The media equation - how people treat computers, TV, and new media like real people and places*. Cambridge University Press (1996).

174. Reilly, D., Rodgers, M., Argue, R., Nunes, M. Marked-up maps: combining paper maps and electronic information resources, In *Personal and Ubiquitous Computing*, Volume 10 (4), 2006.

175. Rekimoto, J.: Pick-and-drop: a direct manipulation technique for multiple computer environments. In *Proceedings of ACM Symposium on User Interface Software and Technology*, UIST'97, ACM Press (1997).

176. Rekimoto, J.: Sensonomy: intelligence penetrating into the real space, In *Proceedings of the international conference on Intelligent User Interfaces*, IUI '09, ACM Press (2009).

177. Rekimoto, J.: SmartSkin: An Infrastructure for Freehand Manipulation on Interactive Surfaces. In *Proceedings of the International conference on Human factors in computing systems*, CHI'02, ACM Press (2002).

178. Resnick P.: Personalized Filters Yes; Bubbles No, Keynote at International Conference on User Modeling, Adaptation and Personalization, UMAP 2011.

179. Rogers, Y., Hazlewood, W., Marshall, P., Dalton, N. S., Hertrich, S.: Ambient Influence: Can Twinkly Lights Lure and Abstract Representations Trigger Behavioral Change? In *Proceedings of the International Conference on Ubiquitous Computing*, UbiComp'10, Springer Verlag (2010).

180. M. Roman, M., Hess, C., Cerqueira, R., Ranganathan, A., Campbell, R., Nahrstedt. K.: A middleware infrastructure for active spaces. In *IEEE Pervasive Computing*, 1(4), 2002.

181. Röcker, C., Hinske, S., Magerkurth, C.: Intelligent privacy support for large public displays. In *Proceedings of the International Conference on Human-Computer Interaction*, HCII'07, Springer Verlag (2007).

182. Rubin, Z.: *Liking and Loving; an invitation to social psychology*, Thomson Learning (1973).

183. Ruiz, J., Li, Y., Lank, E.: User-defined motion gestures for mobile interaction. In *Proceedings of the International conference on Human factors in computing systems*, CHI'11, ACM Press (2011).

184. Rukzio, E., Leichtenstern, K., Callaghan, V., Holleis, P., Schmidt, A., Chin, J.: An Experimental Comparison of Physical Mobile Interaction Techniques: Touching, Pointing and Scanning. In *Proceedings of the International Conference on Ubiquitous Computing*, UbiComp'06, Springer Verlag (2006).

185. Rukzio, E., Mueller, M., Hardy, R.: Design, Implementation and Evaluation of a Novel Public Display for Pedestrian Navigation: The Rotating Compass. In *Proceedings of the International conference on Human factors in computing systems*, CHI'09, ACM Press (2009).

186. Rukzio, E., Schmidt, A., Hussmann, H.: An Analysis of the Usage of Mobile Phones for Personalized Interactions with Ubiquitous Public Displays. In *Workshop on Ubiquitous Display Environments* (2004).

187. Russell, S.J., Norvig, P.: *Artificial Intelligence a modern approach*. Prentice Hall (2003).

188. Salber, D., Dey, A.K., Abowd, G.D.: The context toolkit. In *Proceedings of international conference on Human factors in computing systems*, CHI '99, ACM Press (1999).

189. Schilit, B., Theimer, M.: Disseminating active map information to mobile hosts. In *IEEE Networks* (1994).

190. Schmidt, A.: Implicit Human Computer Interaction Through Context. In Personal and Ubiquitous Computing 4(2/3), 2000.

191. Schmidt, A., Beigl, M., Gellersen, H. There is more to context than location. In *Computer Graphics Journal*, Volume 23 (6), 1999.

192. Schmidt, A., Gross, T., Billinghurst, M.: Introduction to special issue on context-aware computing. In *Computer Supported Cooperative Work*, Volume 13 (2004).

193. Schmidt-Belz B.: User Trust in Adaptvive Systems, In *GI-Workshop "Adapitivität und Benutzermodellierung in interaktiven Softwaresystemen"* (2005).

194. Schöning, J., Krüger, A., Müller, J. Interaction of Mobile Camera Devices with physical maps. In *Proceedings of International Conference on Pervasive Computing*, Pervasive'06, Springer Verlag (2006).

195. Schrammel, J., Gerdenitsch, C., Weiss, A., Kluckner, P., Tscheligi, M.: FORE-Watch – The Clock That Tells You When to Use: Persuading Users to Align Their Energy Consumption with Green Power Availability. In *Proceedings of European Conference on Ambient Intelligence*, AmI'11. Springer Verlag (2011).

196. Seewoonauth, K., Rukzio, E., Hardy, R., Holleis, P.: Touch & Connect and Touch & Select: Interacting with a Computer by Touching it with a Mobile Phone, In *Proceedings of International Conference on Human-Computer Interaction with Mobile Devices and Services*, Mobile HCI'09, ACM Press (2009).

197. Shoemaker, G.: Supporting Private Information on Public Displays, In *Extended Abstracts of International Conference on Human Factors in Computing Systems*, CHI EA'00, ACM Press (2000).

198. Shotton, J., Fizgibbon, A., Cook, M., Sharp, T., Finocchio, M., Moore, R., Kipman, A., Blake, A.: Real-time human pose recognition in parts from single depth images. In *IEEE Computer Vision and Pattern Recognition* (CVPR), IEEE Press (2011).

199. Sixth Sense: http://www.pranavmistry.com/projects/sixthsense/

200. Skog, T., Ljungblad, S., Holmquist, L.: Between Aesthetics and Utility: Designing Ambient Information Visualizations, In *Proceedings on IEEE Conference on Information Visualization*, IEEE Press (2003).

201. Snowdon, D.N., Grasso, A.: Diffusing Information in Organizational Settings: Learning from Experience. In *Proceedings of international Conference on Human factors in computing systems*, CHI'02, ACM Press (2002).

202. Sommer, R.: Personal space: the behavioral basis of design, Prentice-Hall (1969).

203. Sylvester, A., Döring, T., Schmidt, A.: Liquids, Smoke, and Soap Bubbles – Reflections on Materials for Ephemeral User Interfaces, In *Proceedings of the international conference on Tangible, embedded, and embodied interaction*, TEI'10, ACM Press (2010).

204. Taylor, N., Cheverst, K.: Social interaction around a rural community photo display. In *International Journal on Human-Computer Studies*, 67(12), 2009.

205. Tazari, M.: An open distributed framework for adaptive user interaction in ambient intelligence. In *Proceedings of European Conference on Ambient Intelligence*, AmI'10, Springer Verlag (2010).

206. Tschannen-Moran, M., Hoy W.: A multidisciplinary analysis of the nature, meaning, and measurement of trust. In *Review of Educational Research*, Volume 70 (4), 2000.

207. Tufte, E.R.: *The visual display of quantitative information*. Graphics Press (1986).

208. Vertegaal, R., Shell, J.: Attentive user interfaces: the surveillance and sousveillance of gaze-aware objects. In *Social Science Information*, Sur Les Sciences Sociales (2008).

209. Villar, N., Kortuem, G., Van Laerhoven, K., Schmidt, A.: The Pendle: A Wearable Mediator for Mixed Initiative Environments. In *International Workshop on Intelligent Environments*, IEEE Press (2005).

210. Villar, N., Schmidt, A., Kortuem, G., Gellersen, H.: Interacting with proactive public displays, In *Computers and Graphics*, 27(6), Elsevier Science Ltd. (2003).

211. Vogel, D., Balakrishnan, R.: Interactive Public Ambient Displays: Transitioning from Implicit to Explicit, Public to Personal, Interaction with Multiple Users, In *Proceedings of ACM Symposium on User Interface Software and Technology*, UIST'04, ACM Press (2004).

212. Want, R., Fishkin, K., Gujar, A., Harrison, B.: Bridging Physical and Virtual Worlds with Electronic Tags, In *Proceedings of international conference on Human factors in computing systems*, CHI '99, ACM Press (1999).

213. Want, R., Hopper, A., Falcão, V., Gibbons, J.: The Active Badge location system, In *ACM Transactions on Information Systems*, 10(1), 1992.

214. Wagner, J., Lingenfelser, F., André E.: The Social Signal Interpretation Framework (SSI) for Real Time Signal Processing and Recognition. In *Proceedings of Annual Conference of the International Speech Communication Association*, Springer Verlag (2011).

215. Weiser, M.: The computer for the 21st century. In *Scientific American*, September, 1991.

216. Weiser, M., Brown, J.S.: *Designing Calm technology*, Xerox PARC (1995).

217. Wobbrock, J.O., Morris, M.R., Wilson, A.D.: User-defined gestures for surface computing. In *Proceedings of international conference on Human factors in computing systems*, CHI '09, ACM Press (2009).

218. Wu, M., Balakrishnan, R.: Multi-Finger and Whole Hand Gestural Interaction Techniques for Multi-User Tabletop Displays. In *Proceedings of ACM Symposium on User Interface Software and Technology*, UIST'03, ACM Press (2003).

219. Yatani, K., Tamura, K., Hiroki, K., Sugimoto, M., Hashizume, H.: Toss-it: intuitive information transfer techniques for mobile devices. In *Extended Abstracts of international conference on Human factors in computing systems*, CHI EA'05, ACM Press (2005).

220. Zephyr Technology, http://www.zephyr-technology.com/